R E V I E W

On the analysis of non-linear allometries

R O B E R T J . K N E L L   School of Biological and Chemical Sciences, Queen Mary, University of London, U.K.

Abstract. 1. Non-linear allometries are those where a log–log scatterplot of trait size against body size deviates from simple linearity. These are found in many insects, including the horns of beetles, the forceps of earwigs, and the heads of certain castes of ant.

2. Non-linear allometries are often associated with polyphenism that is itself related to behaviour: for example, the alternative mating tactics displayed by many species of beetle are widely associated with dimorphisms in horn size.

3. This paper critically reviews the current techniques used to analyse these datasets.

4. Recommendations include the use of scatterplots and assessment of the goodness of fit of simple linear models as an initial screen for non-linear allometry. The use of recently developed algorithms for 'segmented' regression to analyse continuous allometric relationships, and a pragmatic approach to the analysis of discontinuous relationships that recognises that there is no simple way to distinguish between morphs in some cases, and that all of the proposed methods for doing so have some drawbacks.

5. Worked examples of the analysis of two sets of data from animals that have been the subject of controversy regarding the nature of their allometric relationships are given: further worked examples are provided as online Supporting Information.

Key words. Allometry, *Allomyrina*, analysis, dimorphism, non-linear allometry, polyphenism, switchpoint, *Onthophagus*.

Introduction

The relationship between the size of a particular body part or organ and the size of the whole organism bearing that character is an important and informative aspect of an organism's biology, and such relationships have been studied for many years (Huxley, 1932; D'Arcy Thompson, 1942). This relationship is usually referred to as the *scaling relationship* (Emlen & Nijhout, 2000) or the *static allometry* (Huxley, 1932; D'Arcy Thompson, 1942; Emlen & Nijhout, 2000) of that body part. The majority of such relationships are straight lines, but in the holometabolous insects a number have been described that deviate from simple linearity (Fig. 1), a phenomenon that I shall refer to as *non-linear allometry* [also referred to as *complex allometry* (Nijhout & Wheeler, 1996)]. Complex allometries are of interest to biologists because they can tell us important things about the biology of the species in question. Distributions of data points that can be divided into more than one distinct group, for example, can indicate polyphenism, whereby individuals develop into two or more morphs depending on the environment they experience or their genetic makeup (Eberhard & Gutierrez, 1991;

Correspondence: Robert J. Knell, School of Biological and Chemical Sciences, Queen Mary, University of London, Mile End Road, London E1 4NS, U.K. E-mail: r.knell@qmul.ac.uk
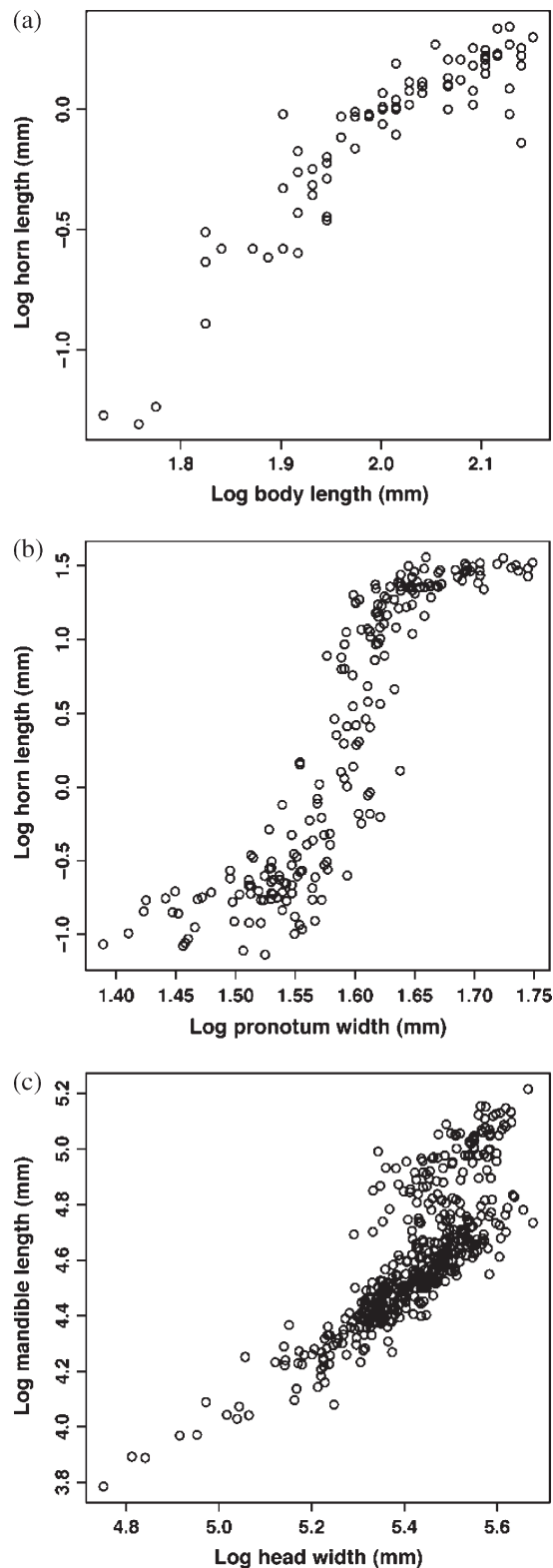
Emlen & Nijhout, 2000). Other possible phenomena that can be indicated by deviations from simple linearity in static allometries are competition for limited resources within the pupae of metamorphosing holometabolous insects (Nijhout & Wheeler, 1996; Knell *et al.*, 2004) and changes in the nature of selection on body parts with increasing body size (Pomfret & Knell, 2006).

In some cases, it may be possible to assign individuals to morphs *a priori*, in which case there are no problems with analysing allometric data, and standard techniques can be followed. Usually, however, there is some degree of uncertainty regarding allocation to morphs or even as to whether there is a dimorphism at all. Typically, statistical analysis is applied to these data sets in order to answer one or more of the following questions: (i) Is there evidence of polyphenism, or is the scaling relationship simply curved or even indistinguishable from a straight line? (ii) If there is polyphenism, is it possible to identify morphological switchpoints (usually a threshold body size) at which animals switch from one morph to another? (iii) Which individuals should be classified as members of which morph? and (iv) Do switchpoints differ between populations of the same species? This is usually done using analyses suggested either by Eberhard and Gutierrez (1991) or Kotiaho and Tomkins (2001), although other approaches have been used as well (Rowland *et al.*, 2005; Cook & Bean, 2006). Whereas the analysis of what might be termed *simple* allometries has been discussed at length

(Warton *et al.*, 2006), there has been very little critical examination of the techniques used to analyse non-linear allometries. The present paper examines the use of these techniques for the detection and further analysis of non-linear allometries, with particular emphasis on dimorphic allometries, and makes recommendations for the best way to analyse a variety of different sorts of such non-linear allometric relationships. The use of these analyses is illustrated by two worked examples, and three further worked examples are provided as online Supporting Information.

Kinds of non-linear allometry

Figure 1 shows scatterplots of non-linear allometries from three insect species. It is obvious from a glance at this figure that there is substantial variation in these relationships, and the form of the relationship is an important consideration in deciding which analysis to use. The simplest and most important division is into what can be called continuous and discontinuous allometries. Continuous allometries are those where the allometric relationship can be regarded as a single, albeit not necessarily straight, line. These allometries can include simple curved relationships (Fig. 1a), ones with a switchpoint where the slope of the relationship changes abruptly [as is claimed, for example, for the horned beetle *Onthophagus binodis* (Cook, 1987; see also Supporting Information Example 1) and for the fig wasp *Sycoscapter australis* (Bean & Cook, 2001)] and also those that show sigmoid patterns such as *Onthophagus taurus* (Fig. 1b). Discontinuous allometries are those that are divided into two (or more) discontinuous groups, which may be more or less separated from each other and which are not adequately modelled by a continuous line. Such discontinuous allometries are described from many dynastid beetles (Emlen & Nijhout, 2000), from a fig wasp (Cook & Bean, 2006), Fig. 1c, and from the earwig *Forficula auricularia* (i.e. Tomkins, 1999; see also Supporting Information Example 3).

Discontinuous allometries can be further divided according to the extent of overlap in the *X* and *Y* variables between the two groups: some have no overlap, some have overlap in *X* (in other words, both morphs are present at some body sizes) and some have overlap in both *X* and *Y*, as in Fig. 1c. It must be remembered that these divisions are not absolute and some cases are difficult to classify: see worked example 2 for one of these.

Is there a non-linear allometry?

Traditionally, simple allometric relationships are described by the equation $y = bx^k$, where y is the size of the organ or body part, x is the size of the whole organism and the exponent

Fig. 1. A variety of non-linear allometries. (a) A smoothly curvilinear relationship. Log horn length plotted against log body length for the dung beetle *Euoniticellus intermedius*. Data from Pomfret and Knell (2006). (b) A sigmoid relationship. Log horn length plotted against log pronotum width for the dung beetle *Onthophagus taurus* (data provided by D. Emlen). (c) A discontinuous relationship with a reasonable amount of separation between the morphs and overlap between the two morphs in both body size and mandible length. Log mandible length plotted against log head width for an undescribed species of fig wasp (Cook & Bean, 2006).

k describes the way that organ size scales with body size, usually estimated as the slope of a straight line fitted to a log–log plot (Huxley, 1932; Gould, 1966; Simmons & Tomkins, 1996; Knell *et al.*, 2004; Tomkins *et al.*, 2005; Warton *et al.*, 2006). The slope of the line is often estimated using major axis (MA) or standardised major axis (SMA) regression instead of conventional linear regression, because these techniques give more accurate estimates of the slope when there is measurement error associated with both the *X* and *Y* measurements. This is often the case in comparative data for multiple species (Warton *et al.*, 2006). Comparable techniques are not yet available for non-linear allometries, however, so the rest of this paper will focus on techniques based on conventional regression techniques. This is unlikely to present a problem so long as measurement error is relatively small, and if a particular dataset is likely to have high measurement error, this can be remedied by using average values from multiple, independent measurements (D. Warton, pers. comm.). Thus, in this context the question 'is there a non-linear allometry?' is equivalent to asking 'if organ size is plotted against body size on a log–log plot, is the relationship adequately explained by a simple linear regression?'.

Eberhard and Gutierrez (1991) fitted quadratic models to untransformed datasets as an initial test to detect non-linear allometry, with a statistically significant quadratic term indicating a possible dimorphism. This approach has often been used since (e.g. Eberhard *et al.*, 2000; Hanley, 2001; Hongo, 2003, in press; Tomkins *et al.*, 2005; Harvey & Gange, 2006; Okada *et al.*, 2006). The use of this test is questionable: there is an implied assumption that all non-linear allometries will be better described by a quadratic curve than by a straight line. This is not always the case. Dimorphic allometries that give roughly symmetrical plots around the major axis will not lead to a significant quadratic term, as seen with the data from *Lucanus cervus* analysed in Supporting Information Example 2. Sigmoid allometries can also return a non-significant quadratic term (Miller & Wheeler, 2005).

Rather than relying on a single statistical test, it is better to recall the question as posed above, whether organ size plotted against body size on a log–log plot is adequately explained by a simple linear regression. This reminds us that we are dealing with a very common question raised in data analysis, and that there are standard methods for assessing how good a fit a linear regression gives. The best place to start is undoubtedly by examining a log–log scatterplot visually, possibly with an appropriate non-parametric smoother such as a cubic spline fitted to the data to aid in the visualisation of any patterns. In many cases this will indicate whether there is a non-linear allometry present, especially if the sample size is a reasonable one (see, for example, Fig. 1b,c). One caveat is that changing the relative *X* and *Y* scales on such a plot can emphasise or minimise apparent differences in slope. Ideally they should be the same, but this may not be possible, in which case care should be taken in the interpretation of any apparent changes. If this is a problem, then plots of standardised normal deviates could be used to remove such scale-dependent effects.

If further analysis is required to demonstrate whether or not there is deviation from a straight line, then a linear regression should be fitted to the data and the standard diagnostic procedures used, in particular an examination of plots of residual versus fitted values. If a satisfactory fit cannot be gained following this procedure, then it will be necessary to identify a suitable model to fit to the data that might describe the patterns better than a simple linear model, and to compare the goodness of fit of both models using standard methods, as discussed in the next section.

Examination of a log–log scatterplot will often also allow a decision to be made regarding whether analysis methods suitable for a continuous or a discontinuous relationship should be used. There will be cases where this is not clear, and in these cases both approaches should be used and final models compared to see which gives the best and most parsimonious fit.

Analysis of continuous relationships

An important question for biologists analysing continuous allometries that are not well described by a straight line, is whether the relationship between organ size and body size is smoothly curvilinear, or whether it is in fact more accurately described by two straight lines that meet at a *switchpoint*. This question is important, because curvilinear relationships can be predicted to arise from competition for resources between rapidly growing body parts in the metamorphosing insect (Nijhout & Wheeler, 1996; Knell *et al.*, 2004), whereas the presence of a switchpoint could indicate some physiological change in the control of development that depends on body size. This could possibly be a *reprogramming* event (Nijhout & Wheeler, 1996).

The model for analysing continuous relationships described in Eberhard and Gutierrez (1991) assumes that the relationship is best described by two straight lines with a single switchpoint. It uses a simple numerical algorithm to find the switchpoint that leads to the highest r^2 value. This procedure is unsatisfactory for two reasons. Firstly, it does not produce confidence intervals for the estimated switchpoint, and secondly, because the switchpoint is estimated before the final model is fitted, incorrect values for the degrees of freedom and for the Akaike information criterion (AIC; see below) are returned. The question of how to identify points in datasets where the response of the *y*-variable to one or more *x*-variables changes suddenly, is one that has been the subject of some interest from statisticians, who usually refer to the problem as *breakpoint*, *segmented* or *change-point* regression (Kim & Siegmund, 1989; Muggeo, 2003). A variety of alternative techniques and software are available to carry out such analyses, for example Segcurve (Luwel *et al.*, 2001), the Segmented package for R (Muggeo, 2003) and Segreg (Oosterbaan, 1994), available from http://www.waterlog.info/segreg.htm.

The estimated switchpoints produced by these alternatives are usually very similar. For the data shown in Fig. 2, for example, both Segmented[1] and the Eberhard and Gutierrez method as implemented in the *switchpoint* function for R (originally written by Ken Wilson of Lancaster University and available from the author) gave identical estimates of the switchpoint up to the fourth significant figure. Segmented (Muggeo, 2003) treats the switchpoint as a true parameter, estimates a standard error for the switchpoint and fits a model with the correct number of degrees of freedom, and it is recommended that this package, or an equivalent be used.
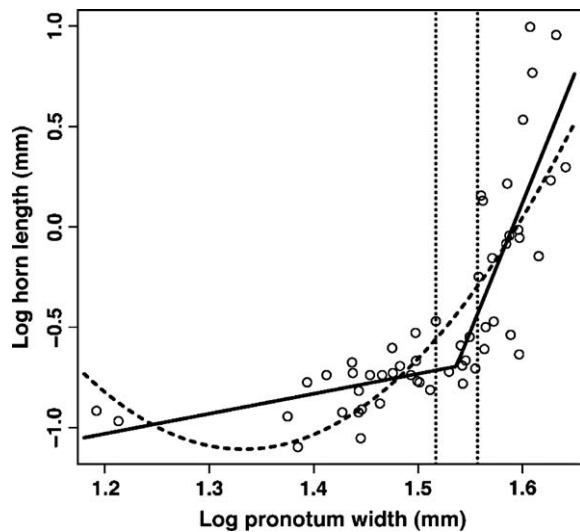
Fig. 2. Log horn length plotted against log pronotum width for the dung beetle *Onthophagus taurus*, originally published in Moczek (2006). The solid line shows a fitted switchpoint model and the dashed line a quadratic model (the power model referred to in the text is not shown for clarity). The vertical dotted lines indicate 95% confidence intervals for the switchpoint. See worked example 1 for details of the analysis.

Once a model with a switchpoint has been fitted to the data, its explanatory power can be compared with appropriate curved or simple linear models. Many solutions to the problem of choosing between different statistical models have been suggested: see, for example Burnham and Anderson (2002); Johnson and Omland (2004) and Olden and Jackson (2000). The examples analysed here use the AIC [calculated as $2 \times$ ($-$log likelihood) $+ 2 \times$ the number of parameters in the model] for model comparison. Models with lower AIC scores are better descriptors of the relationships between the variables, and when comparing two models a difference between AIC scores of more than two is considered to be evidence of a clear difference between them. Models with AIC scores that differ by less than this value are considered to be indistinguishable from each other (Burnham & Anderson, 2002). Note that this is not the only method that can be used to select the most appropriate model. For example, if significance tests are required then fitted models can be compared using partial *F*-tests, or other techniques such as cross-validation (Shao, 1993) could be used. Whatever the technique used, identifying the model that best predicts the data will allow the questions 'Do my data show a dimorphism?' and if so, 'Where is the switchpoint?' to be addressed, and if a breakpoint regression is the best model, the standard errors produced by the breakpoint regression will give an indication of how good the estimate of the switchpoint is.

The next aspect of the analysis to look at is classification into different morphs. For continuous relationships, individuals can be classified into different morphs by asking whether they are larger or smaller than the body-size threshold. The standard errors (or the 95% confidence limits given by ± 1.96 SE) of the threshold can be used to indicate individuals for whom the classification into one or the other morph is uncertain.

One problem with classifying individuals into morphs on the basis of their body size arises when the slope of the allometric relationship is especially steep above the switchpoint. In the case of a dimorphism such as that seen in Fig. 1b, this can lead to some individuals with larger traits being classified as *minors* and some with small traits being classified as *majors*, or vice versa. For some body sizes there is a very wide range of corresponding horn sizes. In an attempt to remedy this, an alternative approach for classifying individuals into separate morphs was proposed by Kotiaho and Tomkins (2001). In this procedure, a switchpoint is estimated based on a regression not of morphological character on body size (as in the other analyses discussed here) but of body size on morphological character. Division of individuals into morphs is therefore based on character size rather than body size. Simulation results suggest, however, that the position of the switchpoint estimated by this method is sensitive to small stochastic variations in the dataset (Fig. 3), indicating that this analysis is not a reliable one for this purpose. There are also problems associated with the assumptions behind regression that arise from regressing *X* on *Y* rather than *Y* on *X*: the two are very different things. If this situation arises, in the absence of behavioural or physiological evidence indicating where the difference between the two morphs lies, the best approach may be to classify those individuals within the 95% confidence intervals as being intermediate between the two morphs.

The final question often asked in studies of dimorphic allometries is whether two or more populations differ in the position of the threshold value. This can be ascertained in one of two ways. Firstly, a *t*-value can be calculated from the estimated switchpoints and standard errors given by Segmented, and then compared with a *t* distribution on $(n_1 - 4) + (n_2 - 4)$ d.f., where n_1 and n_2 are the sample sizes of the two populations. This has to be treated with caution (V. Muggeo, pers. comm.) because the standard errors are only reliable for large sample sizes and/or clear cut relationships. The relevant sampling distribution can vary with the location of the breakpoint and a number of other variables. An alternative is to calculate bootstrap 95% confidence intervals for the difference between the two switchpoints, and to conclude that there is a significant difference between the two if the 95% CIs do not overlap zero. In practice, if Segmented is used for this the algorithm fails to converge in some of the bootstrap replicates, meaning that the results may be unreliable. Since the problems with the Eberhard and Gutierrez method do not affect the estimate of the switchpoint, it is advisable to use this method for this particular analysis.

Sigmoid allometries

Many continuous non-linear allometries are sigmoid, as in Fig. 1b. These are usually regarded as being composed of two elements. Firstly, the part of the curve where the slope increases is thought to be either a breakpoint relationship or an exponential increase in trait size with body size (Tomkins *et al.*, 2005, 2006; Moczek, 2006) (Fig. 2). Secondly, the part of the curve where the slope decreases back towards zero as body size increases [these are the animals described as *asymptotic* majors
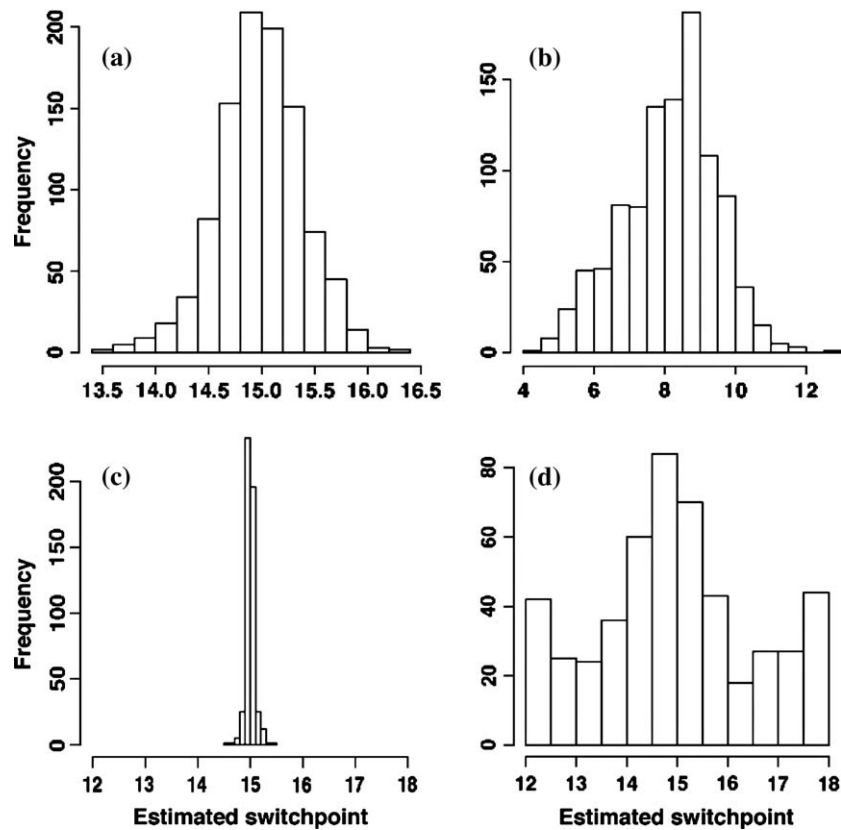
Fig. 3. A comparison of switchpoints estimated using the Eberhard and Gutierrez method and the Kotiaho and Tomkins method. (a) and (b) show results from a simulation of continuous allometries, whereas (c) and (d) are derived from a simulation of discontinuous allometries. For (a) and (b), 1000 simulated bivariate datasets were generated with an *x*-variable mean of 15 and standard deviation of 4. The slope of the *y*-variable was 0.5, with an intercept of zero, unless the *x*-variable was greater than 15, in which case the slope was 1.5. Normally distributed random noise with a standard deviation of 1 was added to the *y*-variable. This gave both *x* and *y* variables with approximately equal ranges. Thus, the true switchpoint for the *x*-variable was 15, and the true switchpoint for the *y*-variable was 7.5. (a) shows the frequency distribution of estimated switchpoints using the Eberhard and Gutierrez method. The distribution is centred around the true switchpoint and nearly 98% of the estimated switchpoints lie within one unit of this value. (b) shows the frequency distribution of the switchpoints estimated using the Kotiaho and Tomkins method. The range of estimates is much wider, with only 44% of them lying within one unit of the true switchpoint. Furthermore, there is a general tendency to overestimate the switchpoint, with the mean estimate being 8.35. Further simulations indicate that this pattern is robust to changes in the position of the switchpoint or to changes in the difference between the slopes above and below the switchpoint, including the case where the slope above the switchpoint is less than that below. Note that the switchpoints were calculated using the numerical search method described by Eberhard and Gutierrez because in most cases the Segmented algorithm failed to converge when calculating switchpoints for the Kotiaho and Tomkins approach. For the simulation of discontinuous data, 1000 datasets were once again generated. The mean value of the *X* variable was 15, with a standard deviation of 3. Below the switchpoint, which was set at 15, the *Y* variable was generated as being equal to the *X* variable (i.e. a slope of 1 and an intercept of zero), with normally distributed noise with a standard deviation of 1 added. Above the switchpoint the *Y* variable was calculated as 1.5 *X* (the slope of the relationship is greater above the switchpoint) plus 3 (which generates a discontinuous relationship with a reasonable amount of separation) plus normally distributed random noise with a standard deviation of 1. Thus the true switchpoint for both variables was 15, although a small amount of increased variation in *Y* should be expected because of the random error added to the datapoints. Note that the estimates in both cases were constrained to be between 12 and 18 in order to allow the large number of calculations to be carried out, because these procedures occasionally produce errors when switchpoints a long way from the centre of the dataset are calculated. (a) shows the distribution of switchpoints calculated using the Eberhard and Gutierrez procedure for discontinuous data, and (b) shows the distribution of switchpoints generated using the Kotiaho and Tomkins approach. The difference between the distributions is striking, and the Kotiaho and Tomkins approach is clearly seriously overestimating or underestimating the switchpoint quite frequently.

in Tomkins *et al.* (2005), although in practice only a few of these curves can be described as asymptotic] may be because of competition for resources in the pupa (Tomkins *et al.*, 2005). However other mechanisms may also contribute to this, such as constraints on horn length or an uncoupling of horn length and body size (Pomfret & Knell, 2006). It is probably best to analyse these components of the curve separately. If an estimate of a switchpoint is required, then only those animals that fall on the lower part of the curve before the slope begins to decrease should be included in the analysis. This can then follow the model comparison procedure outlined above, as in the analysis of the data in Fig. 2.

It is also possible to fit a model to the whole sigmoid curve. Emlen (1996) fitted a modified logistic model to horn length data for *Onthophagus acuminatus* in order to allow residual horn lengths to be calculated for individual beetles. This allowed beetles to be selected on the basis of horn length independent of body size. Moczek *et al.* (2002) fitted a model of the form

$$\text{horn length} = y_0 + \frac{a(\text{body size})^b}{c^b + (\text{body size})^b}$$

to data from *Onthophagus taurus* using non-linear regression. In this case, *c* gives the body size at the lower inflection point on the curve, and this value was used as an indicator of the location of a switchpoint. While this approach may give a reasonable fit to the overall dataset, its use is questionable, firstly because it assumes a smoothly curvilinear transition at the switchpoint, and as we have seen, this is not necessarily the case. Secondly because there is no good reason to assume that these relationships are symmetrical. There are many non-symmetrical curves that could be fitted to such data, but the simplest option in many cases will be to treat the two parts of the curve separately.

Finally, sigmoid allometries lead to markedly bimodal distributions of trait size (Emlen, 1994; Rowland *et al.*, 2005), and an obvious way to separate morphs is by finding the minimum point between the two modes. Eberhard and Gutierrez argue that distinguishing morphs on the basis of frequency distributions is to be avoided, firstly because of the possibility of bias in datasets (collector bias or different rearing conditions, for example, could lead to erroneous bimodal distributions), and secondly because of the statistical difficulty of separating unimodal and bimodal distributions. Nonetheless, in some cases of sigmoid allometry, and especially when body size can be shown to be unimodal when trait size is clearly bimodal, such an analysis can perhaps be justified [see Rowland *et al.* (2005) for a maximum-likelihood approach to doing this]. There are a number of difficulties with this approach, however. Firstly, it provides no information about the location of any switchpoint. Secondly, and more importantly, the location of the second mode in the distribution may well be a consequence of a decline in allometric slope arising from competition for resources within the pupa (Knell *et al.*, 2004). This is a separate process to the differentiation between the morphs, and will strongly affect the location of the minimum point between the two modes. In a system where the decline in slope is weak due to relatively abundant pupal resources, the minimum will be at a greater trait size than in a system where the decline in slope is pronounced.

Worked example 1: horn length in *Onthophagus taurus*

Onthophagus taurus is a dung beetle. Originally from the Mediterranean region, it has been introduced into Australia and America. Large (major) *O. taurus* males carry two curved horns on their heads that they use during combat with conspecific males for access to females. Small (minor) males have reduced horns and use sneak tactics to acquire matings. There has recently been some controversy over the question of whether the dimorphism in these beetles is generated by a developmental

reprogramming event at a certain body size or simply by an exceedingly steep exponential allometric growth of the horns (Tomkins *et al.*, 2005, 2006; Moczek, 2006). Most of the arguments regarding the presence or absence of switchpoints have been made on the basis of visual inspection of scatterplots, however, and none of the authors in this debate have applied rigorous statistical analysis to support their claims. Here, I analyse a dataset of horn length versus pronotum width originally published by Moczek (2006) and shown in Fig. 2. These beetles actually display a sigmoid allometry (as in the other *O. taurus* dataset in Fig. 1b), but here the curve is treated as having two components and all the beetles with a log horn length of greater than 1 were excluded from the analysis. This is because these beetles show a declining slope with increasing body size [the *asymptotic majors* of Tomkins *et al.* (2005)].

A switchpoint model was fitted to the data using the Segmented package (Muggeo, 2003) in R 2.3.1 (The R Foundation for Statistical Computing http://www.R-project.org) and compared with a quadratic model and a power model of the form $y = ax^b$ fitted by the usual means. The switchpoint model gave an estimated switchpoint of 1.537 with a standard error of 0.011, and returned an AIC score of 19.09, substantially less than the quadratic model (26.02) and the power model (68.86). Diagnostic plots revealed no especially influential points that might require closer investigation. On this basis the switchpoint model is the best description of the relationship between horn length and pronotum width in this dataset. This means that this dataset supports the idea that a developmental *reprogramming* event is occurring when the beetles reach a certain size (see the references given above for discussion on what the best horn measurement methodology is and how this can influence whether a switchpoint is found).

A second example of analysis of a continuous relationship (horn size in *Onthophagus binodis*) is given in Supporting Information Example 1.

Discontinuous relationships

If visual inspection of a scatterplot suggests that a discontinuous relationship is present, it is sometimes the case that the discontinuity is sufficiently obvious that further analysis is not needed [as in the case of the dynastid beetle *Chalcosoma atlas* (Emlen, 2001; Kawano, 2002)]. Often, though, further analysis is necessary. In order to answer the question 'do my data show a discontinuous relationship?' it is necessary to fit a model that incorporates the discontinuity and compare its explanatory power with an alternative model or models that do not incorporate a discontinuity. In order to do this, one needs to produce some rule for assigning individuals (i.e. datapoints) into different morphs.

The original analysis suggested for discontinuous allometric data is the procedure outlined by Eberhard and Gutierrez. This involves fitting their model 2 to the data, which takes the form:

$$Y = \beta_0 + \beta_1 X + \beta_2 (X - X^0)D + \beta_3 D + \varepsilon.$$

This assigns individuals into morphs on the basis of a body size switchpoint at X^0. *D* is zero if the *X* value is less than the switchpoint and 1 if it is greater than or equal to it. The switchpoint

is estimated by fitting the model for a variety of values of X^0 and finding the value of X^0 that led to a fit with the lowest value for the adjusted r^2.

Eberhard and Gutierrez suggested that once the switchpoint had been identified, the fitted model could be compared to one that did not include the β_3 term using an *F*-test, with a significantly better fit for the model with the β_3 term indicating the presence of a discontinuity. This approach suffers from the same problems as the corresponding analysis of continuous data, however. The number of degrees of freedom for the fitted models (and their AIC scores) will be incorrect, because the switchpoint is estimated separately and not treated as a real parameter in the model. Furthermore, as noted by the authors, this approach is based on the assumption that overlap in the *x*-variable at the switchpoint is negligible. The simulations shown in Fig. 3 suggest that when this is the case, good estimates of the switchpoint are returned, but both further simulation (not shown) and analyses of real datasets (see Supporting Information Example 2) indicate that if there is overlap in the *x*-variable then this approach tends to identify one end of the region of overlap as the switchpoint. Clearly, if there is overlap in the *x*-variable then the whole concept of a single body size switchpoint is questionable, but even if it is argued that the switchpoint is really the centre of a region of body sizes where animals can develop into either morph, this analysis does not usually identify that point.

One way of approaching the problem of overlap between morphs in the *x*-variable is to use the same model, but with body size as the response variable. Estimating a threshold size for the morphological character is believed to be dimorphic in order to distinguish between morphs (Eberhard *et al.*, 2000; Kotiaho & Tomkins, 2001). This analysis also leads to fitted models with incorrect degrees of freedom and AIC scores, and simulations indicate that, as with continuous relationships, this approach can give unsatisfactory results unless there is no overlap in the *y*-variable and clear separation between the two morphs (Fig. 3).

The two analyses described, thus approach the fundamental problem of assigning individuals into morphs by dividing a bivariate plane into two using either a vertical (Eberhard & Gutierrez, 1991) or a horizontal (Kotiaho & Tomkins, 2001) straight line. An obvious alternative is to use a line with a slope intermediate between these extremes. Cook and Bean (2006) described a discontinuous relationship in male fig wasp mandible allometry with overlap in both the *x*- and the *y*-variable (Fig. 1c). They found both the Eberhard and Gutierrez model and the Kotiaho and Tomkins model to be inadequate for separating the morphs, and separated male fig wasps into morphs based on examination of a frequency histogram of the ratio of mandible length to body size. This was distinctly bimodal, and a separation of morphs based on the point where the two distributions met on the histogram, gave a satisfactory separation of individuals. As the separation of the morphs is based on whether the individual has a ratio of character size to body size that is more or less than a fixed amount, this is equivalent to separating morphs using a straight line through the origin with slope equal to the selected ratio of *y*:*x*.

This distinction worked well in the case of the fig wasps, but this approach again suffers from a number of potential problems. Firstly, the procedure relies on visual inspection of histograms to identify the point of separation between the morphs: this is obviously subjective and may give rather different answers depending on, for example, the number of bins that the data are divided into. This can be avoided by fitting a curve such as a kernel density estimator to the frequency distribution (see Fig. 4b) and identifying the minimum point between the two modes from the fitted curve. Secondly, as with the other analyses, models fitted on the basis of this approach will be calculated with incorrect AIC scores and degrees of freedom. Thirdly, because the line is forced through the origin, the slope will be largely determined by the relative sizes of the measure of the putative dimorphic character and the measure of body size used. If the character in question is relatively small in comparison to the body size measure, the slope will be shallow, and if it is relatively large the slope will be steep, simply as a consequence of where the dataset lies in relation to the origin.

Given the problems arising with this last technique because the line is constrained to pass through the origin, a kind of statistical *reductio ad absurdum* leads to the conclusion that perhaps the best thing to do might be to use a method whereby lines separating the data into two groups are allowed to vary in both slope and intercept. The simple (if inelegant) way to do this is simply to produce a matrix with all possible combinations of intercept and slope between selected intervals between predefined minima and maxima. For each line all the datapoints can be divided into those above and those below the predicted value for that pronotum width and a linear model can be fitted with *morph* as a factor and pronotum width as a continuous variable. The models can be compared on the basis of their goodness of fit (r^2 or AIC: these will give identical rankings in this case). This will find the best line for separating morphs from the predefined subset determined by the intervals and the minima and maxima chosen. In practice, this is computationally intensive and will take a long time for those with slower computers (code to do this in R is available from the author). It can prove difficult to decide which of a multitude of lines giving essentially identical fits should be used and even then the outcome may be no more useful than that from the other methods described here: see Supporting Information Example 2.

A further possibility is to separate individuals into different morphs without a 'rule' of this sort. One way to do this is by the use of finite mixture models. These are a class of statistical models developed for the purpose of analysing datasets that consist of mixtures of data from two or more populations when the provenance of each data point is unknown, by using an *expectation maximisation* algorithm to separate data points into two or more groups. The groups can be distributed along a single axis, as a mixture of two normal distributions (or other forms of frequency distribution), or they can be distributed in more than one dimension. From our point of view, a mixture model that treats the data as a mixture of two linear regressions seems appropriate (McLachlan & Peel, 2000). This approach sounds to be an attractive and general solution to the problem of separating data points in these discontinuous relationships, but in practice fitting mixture models to discontinuous datasets usually seems to lead to a large number of datapoints being misclassified (see the worked examples below). Nonetheless, we should
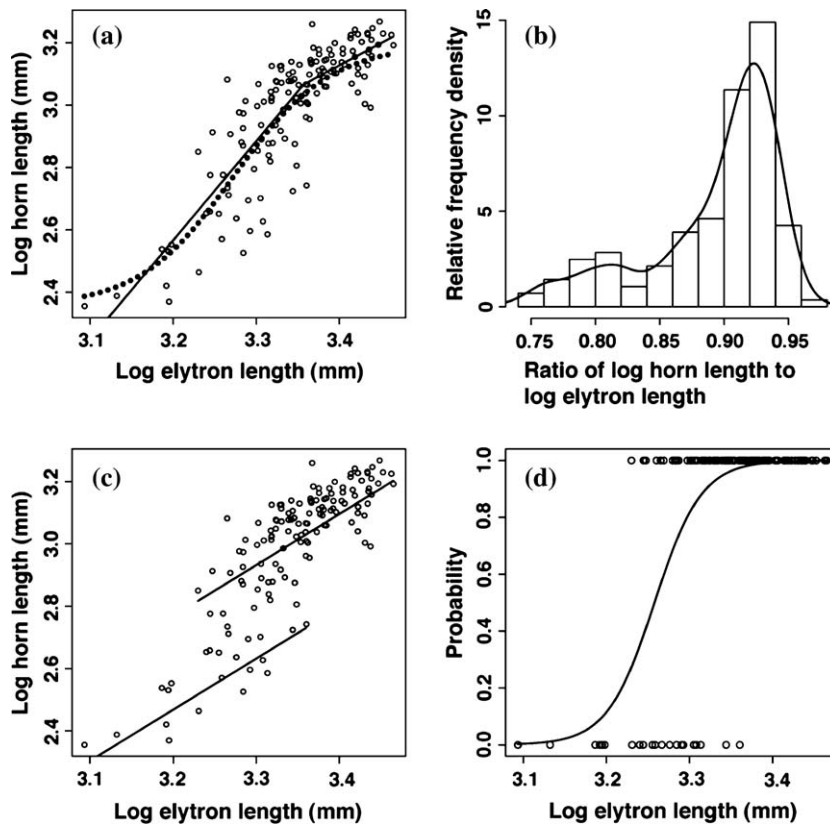
Fig. 4. Horn allometry in *Allomyrina dichotoma*, as discussed in worked example 4. (a) log horn length plotted against log elytra length. The solid line is the fitted breakpoint model from Segmented, and the dotted line is the fitted sigmoid model. (b) Histogram showing the frequency distribution of the ratios of log horn length to log elytron length. The line shows a non-parametric probability density estimate from a kernel density estimator. (c) As (a), but with the fitted discontinuous model with parallel slopes for both morphs. (d) The probability of a male beetle developing into a major morph as determined by fitting a general linear model with binomial errors and a logit link to a dataset with majors coded as 1 and minors coded as zero, with body size as a continuous explanatory variable. The fitted line is $y = e^{-94.9 + 29.8x}/(1 + e^{-94.9 + 29.8x})$.

not discount this approach and it should be considered if other, simpler analysis techniques are not helpful.

An alternative use for finite mixture models arises when using the technique for separating morphs, originally described by Cook and Bean (2006), as discussed above. By treating the set of ratios of trait size : body size as a univariate mixture of two normal (or other) distributions, it is possible to separate morphs using a finite mixture model. This may be helpful in cases where there is no clear bimodality in the frequency distribution. If there is a lot of overlap between the morphs, for example, the distribution of ratios could be skewed or symmetrical with a broad *plateau*. In such cases the mixture model would still allow a objective separation to be made, and the posterior probability calculated for each datapoint would give an indication of confidence for each classification.

Thus, all of the analysis techniques that might be used for separating individuals into one morph or another are flawed in some way, and given the variable nature of these discontinuous allometries there is no one analysis that will be the best in all cases. Given this, a pragmatic approach is called for, with careful visual inspection of scatterplots followed by comparison of the results from the techniques most likely to produce satisfactory results. Once the datapoints have been classified, then it may be necessary to compare the fit of a discontinuous model with a continuous one. If there is negligible overlap between morphs in the *x*-variable, then the Eberhard and Gutierrez (1991) model 2 may be appropriate: otherwise, a general linear model with morph as a factor and the *x*-variable as a continuous variable should be used. This model can then be compared with chosen models with no discontinuity (simple linear, curved or continuous models with a switchpoint, for example), either by means of comparing AIC or by partial *F*-tests. In the former case, the AIC score for the discontinuous model should have 2 added to it to take into account the extra parameter, namely the criterion used to separate the two morphs. In the latter case, the residual degrees of freedom of the discontinuous model should have 1 subtracted before the *F*-test is carried out.

Once it has been decided whether the data are best described by a discontinuous model or not, the remaining description of the data is to quantify the relationship between body size and morph. If there is little or no overlap between the two morphs, then a switchpoint estimated by the Eberhard and Gutierrez procedure might be the best approach. 95% confidence limits can be estimated by bootstrapping if necessary. More commonly, however, there will be overlap in the body sizes that develop into each morph. In this case, the whole concept of a switchpoint of body size becomes questionable. A more suitable approach is to code each datapoint as a 0 or a 1 for minor and major morphs respectively, and to fit a generalised linear model with binomial errors and a logit link to these data, with body size as a continuous explanatory variable. The fitted model will give the probability of developing into a major or a minor morph across the range of body sizes. An alternative is to fit a non-parametric smoother such as a cubic spline to these binomial data (Tomkins & Brown, 2004), but the GLM approach is more informative, since it provides information on both the steepness

of the relationship (the slope of the fitted model) and its location (the easiest measure of this to interpret is the value of the *x*-variable at which 50% of individuals are predicted to develop into each morph). Furthermore, the GLM can easily be extended to allow comparisons between populations or experimental treatments by including these as added factors in the model.

Worked example 2: horn allometry in *Allomyrina dichotoma*

Allomyrina dichotoma (=*Trypoxylus dichotomus*) is a horned beetle that has been the subject of interest recently, from behavioural ecologists interested in subjects including contests between males, the fitness consequences of size and horn length, and the relationship between size and predation risk (Siva-Jothy, 1987; Setsuda *et al.*, 1999; Hongo, 2003, in press; Karino *et al.*, 2004, 2005; Plaistow *et al.*, 2005). A log–log scatterplot of horn length against body size [Fig. 4a, data from Plaistow *et al.* (2005)] clearly shows a non-linear relationship, but it is not obvious whether these data will be best described by a continuous or a discontinuous relationship. Datasets from this beetle have been modelled by different authors as a continuous relationship with a switchpoint (Hongo, 2003, in press), as a continuous sigmoid curve (Karino *et al.*, 2004) and as a discontinuous relationship (Plaistow *et al.*, 2005). In an attempt to resolve this I will compare a continuous breakpoint model, a continuous sigmoidal model, and two discontinuous models on the basis of their AIC scores to determine which gives the best description of the data.

Figure 4a shows the fitted breakpoint model from Segmented and a sigmoid model fitted using the nls (non-linear least squares) function in R. The sigmoid model shown is the one used by Moczek *et al.* (2002), but other sigmoid functions give largely similar fits. Both models describe the data reasonably well towards the larger body sizes, but in the region between a log elytron length of 3.25–3.35 the fit is not so good, with a wide scatter of data points around the line, and the sigmoid curve appears to be overestimating horn length at the smaller body sizes. Figure 4b shows the frequency distribution of the ratios of log horn length to log elytron length to allow morphs to be distinguished using the Cook and Bean approach. The distribution does appear to be bimodal, and fitting a non-parametric probability density estimate with a kernel density estimator enables the minimum point between the two modes to be identified as 0.835. Models were fitted with and without an interaction term between morph and log elytron length, to allow a test of whether the different morphs vary in the slope of the relationship or just in the intercept. The one shown is the model without an interaction. Figure 4c shows the data with the fitted model without an interaction term.

Visual inspection of Fig. 4c suggests that the discontinuous model describes the data better than the two continuous models in Fig. 4a. Although there is a suggestion that horn size at the largest body sizes is being underestimated, this does not appear to be serious when diagnostic plots are examined. This is confirmed by the AIC scores for the fitted models. The breakpoint model scores –207, the sigmoid model –212, the discontinuous model without an interaction term scores –313, and the model with the interaction term –312. Both discontinuous models

therefore give a substantially better description of the data than the continuous models. However, adding the interaction term does not lead to any notable decrease in the AIC score when compared to the model without the interaction term. The information theoretical approach to model selection would then lead to the conclusion that these two models are equally valid descriptions of the data. The alternative approach based on comparing models using *F*-tests (Crawley, 2002) would conclude that the minimal adequate model is the one without an interaction term, since these are not significantly different when compared with a partial *F*-test ($F_{138,137} = 0.6625$, $P = 0.417$). Figure 4d shows the probability of a beetle developing into a minor or a major, calculated by fitting a logistic model to the data on morph coded as 1 for a major and 0 for a minor.

Conclusion

An overriding theme of this paper is that there is no single analysis that is suitable for all forms of non-linear allometry, especially discontinuous allometries. A pragmatic approach is clearly necessary, with careful choice of the most appropriate analysis being made on the basis of the examination of scatterplots and the results of preliminary analysis. The following gives a summary of what I believe to be the key steps in the analysis of a non-linear allometry:

1 Plot a log–log scatterplot of the data, paying attention to the relative axis scaling. Look at the scatterplot and decide whether there is a clear discontinuous relationship, a clear continuous relationship that is not a straight line, a straight line relationship, or whether it is not possible to determine the nature of the relationship from the scatterplot.

2 If the latter, fit a simple linear regression and examine the standard diagnostic plots. If there is no indication of any systematic deviation from a straight line, then the analysis need go no further.

3 If the scatterplot appears to show a continuous relationship that is not well described by a straight line, select some appropriate models that might explain the data well. These can include a breakpoint model, curved models such as a quadratic model and a simple linear model if the results from (2) are inconclusive. Fit the models to the data and compare them on the basis of AIC or *F*-tests. Once the best model (or models) has been identified, then normal model-checking procedures should be followed.

4 If the scatterplot indicates that a discontinuous relationship might exist, then it will be necessary to select a method for distinguishing between morphs. If there is little overlap in the *X*-variable then the Eberhard and Gutierrez model may be used: otherwise, the Cook and Bean approach or possibly a finite mixture model might give good separation. Once a suitable means of separating individual datapoints by morph has been identified, models can be fitted to the data on this basis and compared with other possible models using AIC or *F*-tests, so long as the AIC score or the degrees of freedom are adjusted to account for the extra parameter that has been estimated. If a discontinuous model proves to be the best

description of the data then logistic regression can be used to estimate how the probability of developing into either morph varies with body size.

There are some analyses that could potentially be useful in addressing the questions discussed here, most obviously the use of geometric morphometric techniques to separate morphs. Geometric morphometric techniques [see, for example, Zelditch *et al.* (2004)] are a promising future avenue to follow, assuming that trait shapes as well as sizes change with morph in polyphenic species, but to date, tests of this have not been carried out.

It must always be borne in mind that it may not be possible to separate an allometric relationship into two morphs in a satisfactory way, even when data from other samples of that species clearly show dimorphism. A further point that must be emphasised is that demonstrating a better fit for a model that assumes dimorphism (i.e. a switchpoint model) is not in itself convincing evidence for the existence of, for example, alternative mating tactics or morphs that differ in any other aspect of their behaviour or physiology. Finding a better fit for a model that assumes a dimorphism, shows only that for the models tested the best description of the dataset in question is achieved by that particular model. Especially in cases when the pattern is not particularly clear then these analyses are only the first step towards convincingly demonstrating the existence of a polyphenism, and ideally the analysis of the allometric relationship should be validated with behavioural or physiological data from the species in question. Particular care must be taken with *switchpoints* in continuous datasets. Single datapoints with high leverages can seriously bias these analyses (see Supporting Information Example 1), and possible contributory factors such as changes in shape or changes in landmark orientation might well lead to apparent positive results in these cases when in fact there is no true polyphenism.

Supporting Information

Additional Supporting Information may be found in the online version of this article.

Three worked examples of continuous and discontinuous allometric data from three more insects:

Example 1. The dung beetle *Onthophagus binodis*.

Example 2. The Stag beetle *Lucanus cervus*.

Example 3. The earwing *Forficula auricularia*.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

Notes

[1] The output from Segmented can be interpreted in exactly the same way as the Eberhard and Gutierrez model: the fitted model is essentially the same $[Y = \beta_0 + \beta_1 X + \beta_2 (X - X_0)D]$ with the intercept from segmented $= \beta_0$, the first coefficient $= \beta_1$ and the second coefficient (U.variable.name) $= \beta_2$, and psi $= X_0$.

Acknowledgements

References

Bean, D. & Cook, J.M. (2001) Male mating tactics and lethal combat in the nonpollinating fig wasp *Sycoscapter australis*. *Animal Behaviour*, **62**, 535–542.

Burnham, K.P. & Anderson, D.R. (2002) *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*, 2nd edn. Springer, New York.

Cook, D. (1987) Sexual selection in dung beetles 1. A multivariate study of the morphological variation in two species of *Onthophagus* (Scarabaeidae: Onthophagini). *Australian Journal of Zoology*, **35**, 123–132.

Cook, J.M. & Bean, D. (2006) Cryptic male dimorphism and fighting in a fig wasp. *Animal Behaviour*, **71**, 1095–1101.

Crawley, M.J. (2002) *Statistical Computing: An Introduction to Data Analysis using S-Plus*. John Wiley and Sons, Chichester, U.K.

D'Arcy Thompson, W. (1942) *Growth and Form*. Cambridge University Press, Cambridge, U.K.

Eberhard, W.G., Garcia-C, J.M. & Lobo, J. (2000) Size-specific defensive structures in a horned weevil confirm a classic battle plan: avoid fights with larger opponents. *Proceedings of the Royal Society B – Biological Sciences*, **267**, 1129–1134.

Eberhard, W.G. & Gutierrez, E.E. (1991) Male dimorphisms in beetles and earwigs and the question of developmental constraints. *Evolution*, **45**, 18–28.

Emlen, D.J. (1994) Environmental-control of horn length dimorphism in the beetle *Onthophagus acuminatus* (Coleoptera, Scarabaeidae). *Proceedings of the Royal Society of London, Series B – Biological Sciences*, **256**, 131–136.

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. (2001) Costs and the diversification of exaggerated animal structures. *Science*, **291**, 1534–1536.

Emlen, D.J. & Nijhout, H.F. (2000) The development and evolution of exaggerated morphologies in insects. *Annual Review of Entomology*, **45**, 661–708.

Gould, S.J. (1966) Allometry and size in ontogeny and phylogeny. *Biological Reviews of the Cambridge Philosophical Society*, **41**, 587–640.

Hanley, R.S. (2001) Mandibular allometry and male dimorphism in a group of obligately mycophagous beetles (Insecta: Coleoptera: Staphylinidae: Oxyporinae). *Biological Journal of the Linnean Society*, **72**, 451–459.

Harvey, D.J. & Gange, A.C. (2006) Size variation and mating success in the stag beetle, *Lucanus cervus*. *Physiological Entomology*, **31**, 218–226.

Hongo, Y. (2003) Appraising behaviour during male–male interaction in the Japanese horned beetle *Trypoxylus dichotomus septentrionalis* (Kono). *Behaviour*, **140**, 501–517.

Hongo, Y. (2007) Evolution of male dimorphic allometry in a population of the Japanese horned beetle *Trypoxylus dichotomus septentrionalis*. *Behavioral Ecology and Sociobiology*, **62**, 245–253.

Huxley, J.S. (1932) *Problems of Relative Growth*. Methuen, London.

Johnson, J.B. & Omland, K.S. (2004) Model selection in ecology and evolution. *Trends in Ecology and Evolution*, **19**, 101–108.

Karino, K., Niiyama, H. & Chiba, M. (2005) Horn length is the determining factor in the outcomes of escalated fights among male Japanese horned beetles, *Allomyrina dichotoma* L. (Coleoptera: Scarabaeidae). *Journal of Insect Behavior*, **18**, 805–815.

Karino, K., Seki, N. & Chiba, M. (2004) Larval nutritional environment determines adult size in Japanese horned beetles *Allomyrina dichotoma*. *Ecological Research*, **19**, 663–668.

Kawano, K. (2002) Character displacement in giant rhinoceros beetles. *American Naturalist*, **159**, 255–271.

Kim, H.-J. & Siegmund, D. (1989) The likelihood ratio test for a change-point in simple linear regression. *Biometrika*, **76**, 409–423.

Knell, R.J., Pomfret, J.C. & Tomkins, J.L. (2004) The limits of elaboration: curved allometries reveal the constraints on mandible size in stag beetles. *Proceedings of the Royal Society of London Series B – Biological Sciences*, **271**, 523–528.

Kotiaho, J.S. & Tomkins, J.L. (2001) The discrimination of alternative male morphologies. *Behavioral Ecology*, **12**, 553–557.

Luwel, K., Beem, A.L., Onghena, P. & Verschaffel, L. (2001) Using segmented linear regression models with unknown change points to analyze strategy shifts in cognitive tasks. *Behavior Research Methods, Instruments and Computers*, **33**, 470–478.

McLachlan, G. & Peel, D. (2000) *Finite Mixture Models* (Wiley Series in probability and statistics). John Wiley and Sons, New York.

Miller, K.B. & Wheeler, Q.D. (2005) Asymmetrical male mandibular horns and mating behavior in *Agathidium* Panzer (Coleoptera: Leiodidae). *Journal of Natural History*, **39**, 779–792.

Moczek, A.P. (2006) A matter of measurements: challenges and approaches in the comparative analysis of static allometries. *American Naturalist*, **167**, 606–611.

Moczek, A.P., Hunt, J., Emlen, D.J. & Simmons, L.W. (2002) Threshold evolution in exotic populations of a polyphenic beetle. *Evolutionary Ecology Research*, **4**, 587–601.

Muggeo, V.M.R. (2003) Estimating regression models with unknown break-points. *Statistics in Medicine*, **22**, 3055–3071.

Nijhout, H.F. & Wheeler, D.E. (1996) Growth models of complex allometries in holometabolous insects. *American Naturalist*, **148**, 40–56.

Okada, K., Miyanoshita, A. & Miyatake, T. (2006) Intra-sexual dimorphism in male mandibles and male aggressive behavior in the broad-horned flour beetle *Gnatocerus cornutus* (Coleoptera: Tenebrionidae). *Journal of Insect Behavior*, **19**, 457–467.

Olden, J.D. & Jackson, D.A. (2000) Torturing data for the sake of generality: how valid are our regression models? *Ecoscience*, **7**, 501–510.

Oosterbaan, R.J. (1994). Frequency and regression analysis. *Drainage Principles and Applications* (ed. by H. P. Ritzema), pp. 175–224. ILRI, Wageningen, the Netherlands.

Plaistow, S.J., Tsuchida, K., Tsubaki, Y. & Setsuda, K. (2005) The effect of a seasonal time constraint on development time, body size, condition and morph determination in the horned beetle *Allomyrina dichotoma* L. (Coleoptera: Scarabaeidae). *Ecological Entomology*, **30**, 692–699.

Pomfret, J.C. & Knell, R.J. (2006) Sexual selection and horn allometry in the dung beetle *Euoniticellus intermedius*. *Animal Behaviour*, **71**, 567–576.

Rowland, J.M., Qualls, C.R. & Beaudoin-Ollivier, L. (2005) Discrimination of alternative male phenotypes in *Scapanes australis* (Boisduval) (Coleoptera: Scarabaeidae: Dynastinae). *Australian Journal of Entomology*, **44**, 22–28.

Setsuda, K., Tsuchida, K., Watanabe, H., Kaei, Y. & Yamada, Y. (1999) Size dependent predatory pressure in the Japanese horned beetle, *Allomyrina dichotoma* L. (Coleoptera: Scarabaeidae). *Journal of Ethology*, **17**, 73–77.

Shao, J. (1993) Linear model selection by cross-validation. *Journal of the American Statistical Association*, **88**, 486–494.

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

Siva-Jothy, M.T. (1987) Mate securing tactics and the cost of fighting in the Japanese horned beetle *Allomyrina dichotoma* L. (Coleoptera: Scarabaeidae). *Journal of Ethology*, **5**, 165–172.

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. & Brown, G.S. (2004) Population density drives the local evolution of a threshold dimorphism. *Nature*, **431**, 1099–1103.

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

Tomkins, J.L., Kotiaho, J.S. & LeBas, N.R. (2006) Major differences in minor allometries: a reply to Moczek. *American Naturalist*, **167**, 612–618.

Warton, D.I., Wright, I.J., Falster, D.S. & Westoby, M. (2006) Bivariate line-fitting methods for allometry. *Biological Reviews of the Cambridge Philosophical Society*, **81**, 259–291.

Zelditch, M., Swiderski, D., Sheets, D. & Fink, W.L. (2004) *Geometric Morphometrics for Biologists: A Primer*. Academic Press, London.