



COMMENTARIES

Measuring relative investment: a case study of testes investment in species with alternative male reproductive tactics

JOSEPH L. TOMKINS* & LEIGH W. SIMMONS†

*School of Environmental and Evolutionary Biology, University of St Andrews

†Evolutionary Biology Research Group, Department of Zoology, University of Western Australia

(Received 26 June 2001; initial acceptance 22 August 2001;
final acceptance 13 December 2001; MS. number: SC-1212R)

Evolutionary biologists frequently want to quantify the investment made in a particular trait by individuals or groups. When the trait of interest is linearly related to body size the relative investment of individuals in the population can be described by the allometric relationship between the log-transformed variables (Gould 1966; Prothero 1986; Reiss 1989). The situation becomes more complex when researchers wish to compare the investment in a trait made by different size classes or determine whether the allometric relationships underlying trait expression are the same or different for two or more groups. To draw reliable conclusions, one must account for allometry in an appropriate manner. Here we present a case study from our field of sperm competition and alternative male reproductive tactics. However, the problems and the logic behind the solutions are likely to be the same in many other fields.

Alternative reproductive tactics of males within a species (Gross 1996) are likely to give rise to situations in which there is an asymmetry in sperm competition risk among conspecific males (Parker 1990). This asymmetry arises as a consequence of the behavioural roles that characterize the different male tactics (Parker 1990). Males that guard females tend to be able to prevent rival males from copulating with their partner (Thornhill & Alcock 1983) and in general tend to face a low risk of sperm competition (Parker 1990). In contrast, males that adopt a sneak tactic face a high risk of sperm competition because the female is likely to be subject to previous and/or subsequent matings by the guarding male (Parker 1990). Parker's (1990) models of sperm competition in systems

with alternative male reproductive tactics predict that males adopting the role with greater risk of sperm competition (sneaks) will be selected to evolve adaptations to sperm competition such as the production of larger ejaculates. This prediction has been tested in mammals (Stockley & Purvis 1993), a number of fish (Gage et al. 1995; Taborsky 1998), two species of dung beetle (Simmons et al. 1999) and a ground-nesting bee (Simmons et al. 2000). Selection on testes investment has also been investigated by direct manipulations of sperm competition risk in selected lines of *Drosophila melanogaster* (Pitnick et al. 2001) and yellow dung flies, *Scatophaga (Scathophaga) stercoraria* (Hosken & Ward 2001). There are a number of problems associated with the majority of these studies because testes allometry has not been accounted for correctly. The main points that we make in this paper have been made elsewhere in other contexts, but we show that these have frequently not been taken on board and that the use of the wrong methods can lead to false conclusions.

The methods that have been used to establish differential investment in testes in males with alternative tactics include the gonosomatic index or GSI (the ratio of testes mass to body mass; Gage et al. 1995; Taborsky 1998), residuals from a regression across two tactics (Simmons et al. 1999) or selected lines (Pitnick et al. 2001) and analysis of covariance (Simmons et al. 2000; Hosken & Ward 2001). Here we compare the utility of these methods by reanalysing data from five species that have male dimorphisms associated with alternative reproductive tactics. We show that the methods used are rarely appropriate and suggest a protocol for future studies of testes investment in species with alternative reproductive tactics. We consider 'investment' in a life history sense where individuals trade increased testes size against some other life history variable. The degree to which testes mass deviates positively (high investment) or negatively

Correspondence: J. L. Tomkins, School of Environmental and Evolutionary Biology, Bute Medical Building, University of St Andrews, St Andrews, Fife KY16 9TS, U.K. (email: jtomkins@st-andrews.ac.uk). L. W. Simmons is at the Department of Zoology, University of Western Australia, Nedlands, WA 6009, Australia.

(low investment) from the reference (intraspecific or interspecific) allometry is used as a measure of investment.

The Data

We use both published and unpublished data on testes size variation across alternative male morphologies. Two species of dung beetle, *Onthophagus taurus* and *O. binodis*, have large horned males that guard females and small hornless males that sneak copulations (Simmons et al. 1999). Dawson's burrowing bee, *Amegilla dawsoni*, has a male body size dimorphism where large major males fight for emerging females and small minor males mate opportunistically with foraging females (Alcock 1996); data on testes size come from Simmons et al. (2000). The European earwig, *Forficula auricularia* has large, macro-labic males with long forceps and small or brachylabic males with short forceps (Tomkins 1999), and there is evidence for small males sneaking copulations (J. Tomkins, unpublished data); data have not been published previously and were derived from a population collected from Brownsman Island, Northumberland, U.K. The Atlantic salmon, *Salmo salar*, has large sea-going anadromous males which fight for access to spawning females, and small parr males which remain in their natal streams and sneak copulations; these data come from Gage et al. (1995). Figure 1 shows the relations of log testes mass on log soma mass for each of the data sets.

Analyses and Discussion

Testes allometry

Investment in testes can be analysed in terms of the relationship between log testes mass and log soma mass. For example, allometric exponents greater than one are indicative of greater investment by larger individuals, exponents less than one reveal greater investment by smaller individuals, and isometry indicates equal investment by all individuals. Common use of log-transformed data makes the results of different studies comparable (Prothero 1986). For example, in Table 1 the observed slopes and intercepts (raw data) are influenced by the scale on which the variables were measured, but the log-transformed data are comparable between species and morphs.

A spurious correlation can arise if there is a part-whole relationship between two variables (Sokal & Rohlf 1981; Christians 1999). Body mass is often measured prior to dissection and therefore includes testes mass, which should be subtracted from body mass to yield the mass of the soma. This has not been done in any of the analyses so far reported. Given that organ mass often scales to body mass with an exponent less than one, and small males are often those in the sneak role, the effect of the part-whole correlation will tend to overestimate the investment of sneak males, reducing the allometric exponent. For example, comparing the common least squares regression (LSR) slopes for the small number of species in Table 1 revealed that the slopes of log testes mass on log

body mass tended to be steeper than the slopes of log testes mass on log soma mass, although not significantly so (mean difference 0.049; paired $t_4=2.31$, $P=0.08$).

The use of LSR models for the study of logarithmically transformed variables has been questioned, since least squares techniques assume all of the error to be associated with the estimation of the Y variate (Gould 1966; Sokal & Rohlf 1981; Reiss 1989; Harvey & Pagel 1991). The reduced major axis technique, in contrast, assumes error in the estimation of both X and Y variates and has been recommended for the study of allometry (Gould 1966; Sokal & Rohlf 1981; Reiss 1989; Harvey & Pagel 1991). Nevertheless, where both variates are of the same dimensions, meaningful linear regression equations are achieved using least squares methods (Prothero 1986), and where the correlation coefficient is high ($r>0.95$) there is very little difference between the methods used (Prothero 1986; Reiss 1989). For the dimorphic taxa in Table 1, the correlation coefficients are on the whole low, and the LSR slope substantially underestimates the testes allometry in all but *S. salar*. Prothero (1986) suggested that the most appropriate mathematical fit is determined largely by the desired outcome of the study; evidently if the desire is to predict Y values over the range of observed X values, then least squares techniques are to be recommended.

Finally, establishing the relationship between testes mass and soma mass expected in the absence of sperm competition, effectively the 'null slope', represents a difficulty in interpreting testes allometry (see also Smith 1984). Hosken & Ward's (2001) finding that lines of the yellow dung fly subject to enforced monogamy had lower allometric slopes of testes mass on body mass than polygynous controls is evidence that testes allometry can respond to selection under sperm competition. Increased investment by small males adopting sneak behaviour may well result in a decreased allometric slope across alternative tactics (Stockley & Purvis 1993; Simmons et al. 1999); an allometric exponent less than one indicates that smaller individuals do have relatively larger testes. However, this may simply reflect the existence of an optimal level of investment in testes that maximizes fertility. All individuals will reach the optimum that will constitute a larger proportion of the soma for small males. This can be true in the absence of sperm competition (Simmons et al. 2000). Thus, within species or groups, studies of the relative influence of natural selection versus sperm competition on testes investment across males of different size classes can be confounded.

Gonosomatic index

Gonosomatic index (GSI) is an intuitively appealing way of measuring testes investment because it is believed that variation in testes weight that is due to body weight can be controlled by dividing by body weight. Unfortunately there are a number of underlying assumptions to this method that are not readily apparent (Pearson 1896; Tanner 1949; Atchley 1978; Atchley & Anderson 1978; de Vlaming et al. 1982; Prothero 1986; Packard & Boardman 1987, 1989, 1999; Raubenheimer 1995). The primary problem is that ratios do not 'control' for body weight

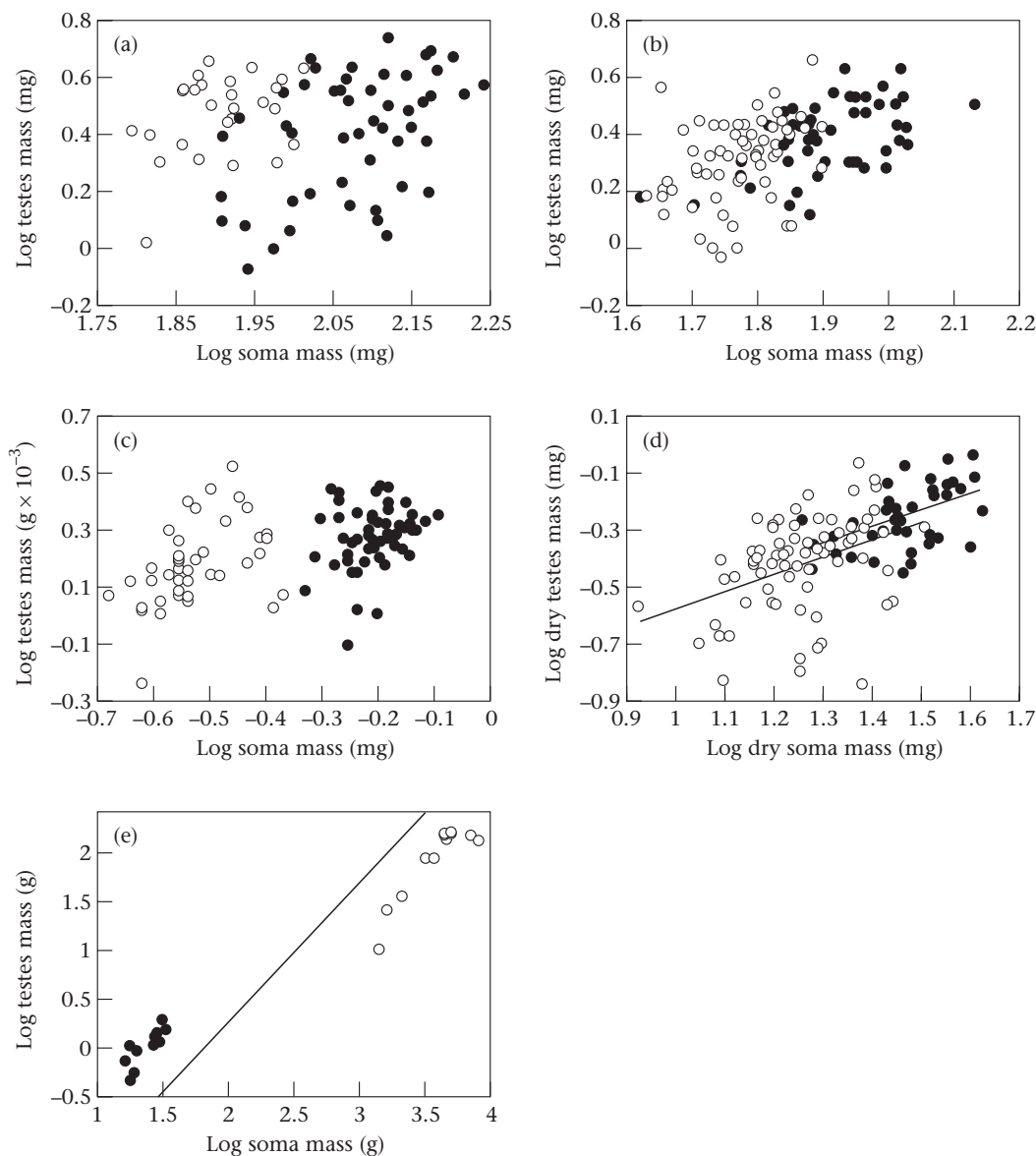


Figure 1. Scatterplots of log testes mass on log soma mass for five dimorphic taxa: (a) *Onthophagus binodis*, (b) *Onthophagus taurus*, (c) *Amegilla dawsoni*, (d) *Forficula auricularia* and (e) *Salmo salar*. ●: Guards in *O. binodis* and *O. taurus*, majors in *A. dawsoni*, macrolabics in *F. auricularia* and anadromous males in *S. salar*. ○: Sneaks in *O. binodis* and *O. taurus*, minors in *A. dawsoni*, brachylabics in *F. auricularia* and parr males in *S. salar*. The LSR regression slopes are shown for *F. auricularia*, and the slope from which least square means are calculated for the ANCOVA is shown for *S. salar*.

unless the trait of interest is isometric (intercept of zero and slope of one). Where, for example, the intercept of testes mass on body mass is less than zero, a positive relationship between GSI and body mass will be generated. In contrast, if the intercept is positive the relationship will be negative (e.g. Figure 1 in Packard & Boardman 1999). Table 1 shows the intercepts for the observed (untransformed) data from which GSI is calculated. In all species except *S. salar*, the intercept of the common slope was significantly greater than zero. In the case of *S. salar* a common regression constructed across morphs would be statistically invalid, owing to the gross differences between groups in both soma mass and testes mass. The

intercepts for both anadromous and parr males were significantly less than zero.

As outlined above, an assumption of testes isometry is particularly pertinent to the study of alternative reproductive tactics and the evolutionary consequences of sperm competition. Small males generally adopt the sneak role because they are unsuccessful in competition for access to females. Unfortunately however, with negative allometry, smaller individuals will have higher GSI (Simmons et al. 1999). Negative allometry is not necessarily predicted solely by sperm competition and is a common scaling relationship of organ mass on body mass (e.g. brain size, Martin 1981). Hence when GSI is used in

Table 1. Regression statistics for five male dimorphic taxa

Species	Observed LSR slope	Observed intercept	Log-log LSR slope	Log-log intercept	Log-log <i>r</i>	RMA slope
<i>O. binodis</i>						
Guard	0.020±0.006**	0.314±0.769	1.048±0.300***	-3.002±0.970**	0.45	2.348
Sneak	0.023±0.014	1.142±1.168	0.816±0.438	-1.860±1.230	0.35	2.340
Both	0.008±0.004	2.022±0.466***	0.289±0.186	-0.459±0.576	0.17	1.643
<i>O. taurus</i>						
Guard	0.030±0.008**	0.304±0.515	0.610±0.231*	-0.769±0.444*	0.37	1.650
Sneak	0.017±0.007*	1.121±0.592	0.804±0.250**	-1.122±0.442**	0.37	2.162
Both	0.021±0.003***	0.874±0.254***	0.683±0.115***	-0.908±0.211***	0.49	1.390
<i>A. dawsoni</i>						
Major	1.310±0.850*	1.136±0.536*	0.543±0.293	0.389±0.062***	0.24	2.242
Minor	3.923±1.522*	0.410±0.470	0.845±0.275**	0.623±0.146***	0.44	1.916
Both	1.238±0.288***	1.196±0.150***	0.365±0.073***	0.360±0.027***	0.45	0.800
<i>F. auricularia</i>						
Macrolabic	0.013±0.004**	0.192±0.118	0.602±0.196**	-1.144±0.292***	0.47	1.311
Brachylabic	0.013±0.003***	0.172±0.064**	0.419±0.136**	-1.171±0.206***	0.36	1.292
Both	0.013±0.001***	0.159±0.039***	0.576±0.084***	-1.266±0.121***	0.56	1.077
<i>S. salar</i>						
Anadromous	0.021±0.005**	16.076±24.630	1.485±0.216***	-3.383±0.774**	0.91	1.635
Parr	0.062±0.014**	-0.281±0.342	1.286±0.338**	-1.710±0.454**	0.78	1.638

Observed least squares regression (LSR) slope and intercept, and testes allometry examined by LSR, correlation coefficient (*r*) and reduced major axis (RMA) are shown±SE. All values are calculated with soma mass (body mass–testes mass) to control for the part–whole correlation that arises when body mass is used.

P*<0.05; *P*<0.01; ****P*<0.001.

Table 2. Gonosomatic index (GSI), residual testes mass and adjusted testes mass±SE from ANCOVA for five male dimorphic taxa

Species	GSI	<i>t</i> _{GSI}	Residual testes mass	<i>t</i> _{residual}	Adjusted testes mass†
<i>O. binodis</i>					
Guard	0.022±0.001	6.54***	-0.039±0.025	2.989**	2.186±1.066
Sneak	0.036±0.001		0.075±0.034		3.784±1.104
<i>O. taurus</i>					
Guard	0.0357±0.001	2.414*	-0.003±0.018	0.179	2.173±1.057
Sneak	0.0311±0.001		0.002±0.015		2.218±1.045
<i>A. dawsoni</i>					
Major	0.003±0.0001	8.67***	-0.008±0.016	0.723	1.513±1.074
Minor	0.005±0.0002		0.001±0.018		2.065±1.099
<i>F. auricularia</i>					
Macrolabic	0.019±0.7 ⁻³	2.321*	0.011±0.022	0.686	0.425±1.042
Brachylabic	0.020±0.001		-0.005±0.015		0.458±1.069
<i>S. salar</i>					
Anadromous	0.235±0.002	5.51***	—	—	2.167±1.513
Parr	0.465±0.003		—		49.88±1.588

All values are calculated using soma mass (body mass–testes mass) to control for the part–whole correlation that arises when body mass is used.

†In mg, apart from *S. salar* (g).

P*<0.05; *P*<0.01; ****P*<0.001.

ignorance of the underlying relationship between testes mass and body mass, increased testes investment by small sneak males can be found irrespective of sperm competition.

Table 1 shows the observed relationship between testes mass and soma mass across morphs, from which GSI is calculated. In all species except *A. dawsoni*, the observed slope is substantially less than one. Table 2 shows the GSI values for the five dimorphic species studied; in each case the small males have significantly higher GSI values than

the large males. This relationship is generated because of violations of the assumptions for the use of ratios. Thus the significance of greater GSI in sneaks cannot be separated from what might simply be the underlying scaling relationship of testes mass on body mass. This point is highlighted in *A. dawsoni*, in which the greater GSI of small minor males suggests adaptations to sperm competition; however, female bees are known to be monandrous (Simmons et al. 2000). The use of GSI also assumes a single linear relationship when gonad weight is

Table 3. Analysis of covariance of log testes mass on male tactic, with the covariate log soma mass

Source	Full model				Reduced model			
	df	Mean square	F	P	df	Mean square	F	P
<i>O. binodis</i>								
Model	3	0.193	6.461	0.0006	2	0.286	9.669	0.0002
Tactic	1	0.012	0.389	0.5345	1	0.488	16.528	0.0001
Log soma mass	1	0.285	9.537	0.0028	1	0.498	16.857	0.0001
Tactic*log soma mass	1	0.007	0.241	0.6248				
Error	74	0.029			75	0.029		
<i>O. taurus</i>								
Model	3	0.187	11.734	0.0001	2	0.277	17.555	0.0001
Tactic	1	0.005	0.294	0.5885	1	0.001	0.072	0.7891
Log soma mass	1	0.261	16.343	0.0001	1	0.272	17.141	0.0001
Tactic*log soma mass	1	0.005	0.312	0.5777				
Error	109	0.016			110	0.016		
<i>A. dawsoni</i>								
Model	3	0.146	10.114	0.0001	2	0.215	14.957	0.0001
Tactic	1	0.036	2.486	0.1182	1	0.058	4.022	0.0478
Log soma mass	1	0.169	11.698	0.0009	1	0.199	13.879	0.0003
Tactic*log soma mass	1	0.008	0.564	0.4543				
Error	94	0.014			95	0.014		
<i>F. auricularia</i>								
Model	3	0.357	18.983	0.0001	2	0.535	28.732	0.0001
Tactic	1	0.000	0.004	0.9509	1	0.013	0.713	0.4001
Log soma mass	1	0.276	14.706	0.0002	1	0.401	21.504	0.0001
Tactic*log soma mass	1	0.000	0.000	0.9908				
Error	111	0.019			112	0.019		
<i>S. salar</i>								
Model	3	7.419	329.043	0.0001	2	11.126	514.291	0.0001
Tactic	1	0.082	3.640	0.0716	1	0.264	12.214	0.0023
Log soma mass	1	0.827	36.684	0.0001	1	1.594	73.670	0.0001
Tactic*log soma mass	1	0.004	0.190	0.6676				
Error	19	0.022			20	0.022		

The ANCOVAs are reported before and after the removal of the interaction term.

regressed on body weight (de Vlaming et al. 1982). However, divergent selection on testes investment means that the allometry of testes weight may well differ between groups of individuals, as found in monogamous and polygynous lines of *S. stercoraria* (Hosken & Ward 2001).

Extensive discussions of the use and misuse of ratio data by a number of authors (Pearson 1896; Tanner 1949; Atchley 1978; Atchley & Anderson 1978; de Vlaming et al. 1982; Prothero 1986; Packard & Boardman 1987, 1988, 1999; Raubenheimer 1995; Smith 1999) suggest that the GSI is highly unlikely to be a useful descriptor of gonad investment for any species.

Residual testes mass

Residual analysis has been used to quantify relative testes investment (Simmons et al. 2000; Pitnick et al. 2001). Table 2 shows the residual testes mass taken from linear regressions of log testes mass on log soma mass across alternative tactics. Again, such an analysis would be invalid for *S. salar* because the gross differences in body mass preclude the common regression. Only in *O. binodis* is there a significant difference between the testes investment of large and small males measured using residuals. However, comparison of the residual analysis in Table 2 and the analysis of covariance (ANCOVA) results in Table 3 shows that although some

of the results are qualitatively similar, for *A. dawsoni* residual analysis and ANCOVA yield conflicting results.

The use of residuals from LSR ignores the possibility that the relationship between testes mass and body mass may differ between the two male tactics, something that is addressed by ANCOVA. Residual analysis is therefore not suitable where two or more size classes of individual are involved and should not be used as a method for comparing testes investment in dimorphic species, or selected lines where body size differences exist. For further discussion of the problems of using LSR residuals see Smith (1984), García-Berthou (2001) and Darlington & Smulders (2001).

Analysis of covariance

In the investigation of testes investment, we are interested in testes size as an adaptation to sperm competition, and we wish to conduct an analysis that will therefore detect any divergence from the underlying soma mass–testes mass relationship. ANCOVA does account for this underlying relationship and hence is an appropriate model to detect adaptations to sperm competition that are represented by increased elevation of the testes allometry in one tactic.

When examining the relationship between soma mass and testes mass of two alternative tactics it is important to

establish whether the slopes are homogeneous. Homogeneity of variance is examined within ANCOVA from the interaction term between mating tactic and the covariate log soma mass (Table 3). As with analysis of variance, nonsignificant interaction terms can be removed from the model, pooling this variance with the error variance (Hendrix et al. 1982). However, the ability to detect significant interactions is weak, and to avoid committing type II errors, one should discard interactions only where $P > 0.2$ (Hendrix et al. 1982). An alternative approach to the removal of the nonsignificant interaction is to use a model selection approach (Anderson et al. 2000). This methodology will determine whether the model including the interaction represents the most appropriate fit to the data, based on log-likelihood theory. We have used the method of Hendrix et al. (1982). For all of the species reviewed here the interaction terms were nonsignificant (Table 3) and were removed (Table 3). The removal of the nonsignificant interaction is an important step in exploring relative testes investment, as the omission of this step would have resulted in the conclusion that there were no differences in testes investment in three species in which it was detected in the reduced model. We failed to remove the nonsignificant interaction term in Simmons et al. (2000) and concluded that there was no difference between the investment of minor and major male *A. dawsoni* in their testes.

It is important to realize that even when slopes are homogeneous, residual analysis and ANCOVA are not equivalent, the difference lying in the calculation of the common slope. ANCOVA does not use an LSR slope derived from the pooled data, but rather the common slope b_c is estimated from the ratio of the pooled ANOVA error terms for each treatment (Snedecor & Cochran 1980; Sokal & Rolf 1981). Estimation of the LSR slope would involve fitting a line through the data for alternative male tactics so that the residuals about the line are equally distributed across alternative tactics, ignoring the differences in elevation between tactics, and thus yielding little difference between them. The estimation of b_c removes this problem because the pooled slope calculated from an ANCOVA is not drawn through the scatter for each tactic so that residual values reflect the different elevations of the data. For clarity on this point (it is not convention to show the line) we have included the pooled slope of the ANCOVA in Fig. 1e. For this slope, from which the least squares means are calculated in the ANCOVA, b_c is 1.451, and the intercept is -2.59 , both lying between the slopes and intercepts of the parr and anadromous males (Table 1, Fig. 1e). For further recent discussion of the misuse of residuals in ecology and the superiority of ANCOVA see García-Berthou (2001) and Darlington & Smulders (2001). Pitnick et al. (2001) recently reported divergent residual testes masses for monogamous and polyandrous lines of *D. melanogaster* after 61 generations of selection. The extreme divergence in body size between the lines and the divergent allometries in the similar experiment on *S. stercoraria* (Hosken & Ward 2001) suggest that ANCOVA might have been a more appropriate analysis for these data.

The reduced model (Table 3) assumes that the slopes of the two tactics are equal, and compares the adjusted or least squares mean of each tactic at the mean value of the covariate (Snedecor & Cochran 1980). The least squares means are the mean values of the dependent variable (log testes mass) at the mean value of the covariate (log soma mass) controlling for testes allometry. This analysis reveals significant differences between the testes investment of large and small males in *O. binodis* and *S. salar*, but not in *O. taurus* (Table 3). Fortunately these results are the same as those reported in the original papers (Gage et al. 1995; Simmons et al. 1999). There was no difference between the adjusted testes masses of brachylabic and macrolabic *F. auricularia* (Table 3). There was also a significant difference in the adjusted testes mass of major and minor *A. dawsoni*.

Where the interaction term of the ANCOVA is significant the slope of testes allometry differs between the tactics. Here the interpretation of main effects becomes problematic because the magnitude of the difference between the two tactics will vary according to the value of the covariate at which they are compared (Snedecor & Cochran 1980; Henderson 1982; Hendrix et al. 1982). Thus, unless there is an a priori expectation for which male size class should have the greater testes allometry, interpretation of the slopes is likely to be ambiguous.

Significant heterogeneity of the testes allometry does not occur in the data we review here; however, it is not unlikely that future studies will find such a situation. Snedecor & Cochran (1980) presented a method for estimating the average difference between the treatments taken over one of the populations. In this method the slope of one of the samples only is used in the covariance adjustment. An additional method that is perhaps easier to perform is to use the pooled slope, but compare the difference between the tactics at the mean as well as at 1 SD above and below the mean (Hendrix et al. 1982). Hendrix et al. (1982) achieved this first by adding 1 SD to the covariate of one treatment and comparing the adjusted means and, second, by subtracting 1 SD from the original values of one covariate and then again comparing the adjusted means. If the slopes intersect close to the mean of the covariate (log soma mass), then there may be a difference between the adjusted mean (and adjusted mean ± 1 SD) in one direction but not in the other (e.g. Hendrix et al. 1982). If, however, the slopes intersect far from the mean, the adjusted values both at the mean and at the mean ± 1 SD may be significantly different between the treatments (morphs) and some interpretation of the average difference between the tactics could be made (Snedecor & Cochran 1980).

Limitations

It can be seen from Fig. 1 that there is a great deal of scatter in the plots of testes mass on soma mass. This scatter means that large sample sizes are needed to detect heterogeneity in testes allometry. For example, the analysis of our data for *A. dawsoni* reveals that there is a significant difference in the level of investment of the major and minor males, with minors investing significantly more than major males. This is unexpected in a

system in which we are sure that there is no sperm competition (Simmons et al. 2000). The answer may well be that there is a threshold investment in testes that all males are selected to reach to fertilize a female's eggs successfully, such that the relationship between testes mass and soma mass across the morphs would be asymptotic. The different investment revealed by the ANCOVA may therefore be an artefact of the testes allometry of minor males representing the part of the curve below the asymptote. A larger sample size might have detected a heterogeneity in the slopes of the major and minor males that reflects this allometric trajectory towards the threshold in minors and no relation in majors.

The *S. salar* data illustrate a concern raised by some authors over using the ANCOVA for data sets that do not overlap (Packard & Boardman 1999). This concern arises because the mean size at which the least squares means are generated is well outside the range of values for both the anadromous and parr males. If, however, the ANCOVA is considered, not to construct a hypothetical individual of mean size, but rather as a method of standardizing the body size differences and examining the exponents and elevations of testes mass in each morph, the ANCOVA remains valid.

Conclusion

Unfortunately it appears that all the analyses of relative testes investment in dimorphic taxa and one of the two selection experiments have been in some way flawed. The errors in these analyses represent a significant problem since intraspecific studies, and selection experiments, are otherwise a powerful way of testing sperm competition theory. Fortunately, the majority of the findings from this reanalysis were consistent with the original interpretations of the data and the biology of the species involved. We recommend that testes investment be interpreted in the context of testes allometry using log-transformed data, and that the ANCOVA be used as the first tool in investigating divergent testes investment. Should heterogeneity of slopes be discovered, we recommend that adjusted testes mass be used to estimate testes investment of individuals ± 1 SD of the mean body mass. This methodology has broader significance than testes investment, and will be useful wherever researchers are interested in size-independent measures of investment (e.g. condition, see García-Berthou 2001). We hope that the errors revealed here, many of which we have made in the past, will be avoided by other researchers in the future.

We thank Michael J. Reiss, Janne Kotiaho, Lotta Kvarnemo, Natasha LeBas, John Hunt, David Hosken, Matt Gage and Ben Sheldon for constructive criticism. Matt Gage generously provided us with the salmon data. J. L. T. was funded by a University Postdoctoral Fellowship from the University of Western Australia and a David Phillips Fellowship from the BBSRC. L. W. S. was funded by the Australian Research Council.

References

- Alcock, J. 1996. The relation between male body size, fighting, and mating success in Dawson's burrowing bee, *Amegilla dawsoni* (Apidae, Apinae, Anthophorini). *Journal of the Zoological Society of London*, **239**, 663–675.
- Anderson, D. R., Burnham, K. P. & Thompson, W. L. 2000. Null hypothesis testing: problems, prevalence and an alternative. *Journal of Wildlife Management*, **64**, 912–923.
- Atchley, W. R. 1978. Ratios, regression intercepts and the scaling of data. *Systematic Zoology*, **27**, 78–83.
- Atchley, W. R. & Anderson, D. E. 1978. Ratios and the statistical analysis of biological data. *Systematic Zoology*, **25**, 137–148.
- Christians, J. K. 1999. Controlling for body mass effects: is part-whole correlation important? *Physiological and Biochemical Zoology*, **72**, 250–253.
- Darlington, R. B. & Smulders, T. V. 2001. Problems with residual analysis. *Animal Behaviour*, **62**, 599–602.
- Gage, M. J. G., Stockley, P. & Parker, G. A. 1995. Effects of alternative male mating strategies on characteristics of sperm production in the Atlantic salmon (*Salmo salar*): theoretical and empirical investigations. *Philosophical Transactions of the Royal Society of London, Series B*, **350**, 391–399.
- García-Berthou, E. 2001. On the misuse of residuals in ecology: testing regression residuals vs. the analysis of covariance. *Journal of Animal Ecology*, **70**, 708–711.
- Gould, S. J. 1966. Allometry and size in ontogeny and phylogeny. *Biological Reviews*, **41**, 587–640.
- Gross, M. R. 1996. Alternative reproductive tactics: diversity within sexes. *Trends in Ecology and Evolution*, **11**, 92–98.
- Harvey, P. H. & Pagel, M. D. 1991. *The Comparative Method in Evolutionary Biology*. Oxford: Oxford University Press.
- Henderson, C. R. 1982. Analysis of covariance in the mixed model: higher level, non homogeneous, and random regressions. *Biometrics*, **38**, 623–640.
- Hendrix, L. J., Carter, M. W. & Scott, D. 1982. Covariance analyses with heterogeneity of slopes in fixed models. *Biometrics*, **38**, 641–650.
- Hosken, D. J. & Ward, P. I. 2001. Experimental evidence for testes size evolution via sperm competition. *Ecology Letters*, **4**, 10–13.
- Martin, R. D. 1981. Relative brain size and metabolic rate in terrestrial vertebrates. *Nature*, **293**, 57–60.
- Packard, G. C. & Boardman, T. J. 1987. The misuse of ratios to scale physiological data that vary allometrically with body size. In: *New Directions in Ecological Research* (Ed. by M. E. Feder, A. F. Bennett, W. W. Burggren & R. B. Huey), pp. 216–239. Cambridge: Cambridge University Press.
- Packard, G. C. & Boardman, T. J. 1988. The misuse of ratios, indices and percentages in ecophysiological research. *Physiological Zoology*, **61**, 1–9.
- Packard, G. C. & Boardman, T. J. 1999. The use of percentages and size-specific indices to normalize physiological data for variation in body size: wasted time, wasted effort? *Comparative Biochemistry and Physiology, Part A*, **122**, 37–44.
- Parker, G. A. 1990. Sperm competition games: sneaks and extra-pair copulations. *Proceedings of the Royal Society of London, Series B*, **242**, 127–133.
- Pearson, K. 1896. Mathematical contributions to the theory of evolution. On a form of spurious correlation which may arise when indices are used in the measurement of organs. *Proceedings of the Royal Society of London*, **60**, 489–498.
- Pitnick, S., Miller, G. T., Reagan, J. & Holland, B. 2001. Males' evolutionary responses to experimental removal of sexual selection. *Proceedings of the Royal Society of London, Series B*, **268**, 1071–1080.

- Prothero, J.** 1986. Methodological aspects of scaling in biology. *Journal of Theoretical Biology*, **118**, 259–286.
- Raubenheimer, D.** 1995. Problems with ratio analysis in nutritional studies. *Functional Ecology*, **9**, 21–29.
- Reiss, M. J.** 1989. *The Allometry of Growth and Reproduction*. Cambridge: Cambridge University Press.
- Simmons, L. W., Tomkins, J. L. & Hunt, J. C.** 1999. Sperm competition games played by dimorphic male beetles. *Proceedings of the Royal Society of London B*, **266**, 145–150.
- Simmons, L. W., Tomkins, J. L. & Alcock, J. A.** 2000. Can minor males of Dawson's burrowing bee, *Amegilla dawsoni* (Hymenoptera: Anthophorini) compensate for reduced access to virgin females through sperm competition? *Behavioral Ecology*, **11**, 319–325.
- Smith, R. J.** 1984. Determination of relative size: the 'criterion of subtraction' problem in allometry. *Journal of Theoretical Biology*, **108**, 131–142.
- Smith, R. J.** 1999. Statistics of sexual size dimorphism. *Journal of Human Evolution*, **36**, 423–459.
- Snedecor, G. W. & Cochran, W. G.** 1980. *Statistical Methods*. Ames, Iowa: Iowa State University Press.
- Sokal, R. R. & Rohlf, F. J.** 1981. *Biometry*. New York: W.H. Freeman.
- Stockley, P. & Purvis, A.** 1993. Sperm competition in mammals: a comparative study of male roles and relative investment in sperm production. *Functional Ecology*, **7**, 560–570.
- Taborsky, M.** 1998. Sperm competition in fish 'bourgeois' males and parasitic spawning. *Trends in Evolution and Ecology*, **16**, 222–227.
- Tanner, J. M.** 1949. Fallacy of per-weight and per-surface area standards, and their relation to spurious correlation. *Journal of Applied Physiology*, **2**, 1–15.
- Thornhill, R. & Alcock, J.** 1983. *The Evolution of Insect Mating Systems*. Cambridge, Massachusetts: Harvard University Press.
- 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.
- de Vlaming, V., Grossman, G. & Chapman, F.** 1982. On the use of Gonosomatic index. *Comparative Biochemistry and Physiology*, **73A**, 31–39.