Why are metabolic scaling exponents so controversial? Quantifying variance and testing hypotheses

Abstract

The metabolic theory of ecology links physiology with ecology, and successfully predicts many allometric scaling relationships. In recent years, proponents and critics of metabolic theory have debated vigorously about the scaling of metabolic rate. We show that the controversy arose, in part, because researchers examined the mean exponent separately from the variance. We estimate both quantities simultaneously using linear mixed-effects models and data from 1242 animal species. Metabolic rate scaling converges on the predicted value of 3/4 but is highly heterogeneous: 50% of orders lie outside the range 0.68–0.82. These findings are robust to several forms of statistical uncertainty. We then test competing hypotheses about the variation. Metabolic theory is currently unable to explain differences in scaling among orders, but the patterns are not consistent with competing explanations either. We conclude that current theories are inadequate to explain the full range of metabolic scaling patterns observed in nature.

Keywords

Allometry, body size, energetics, macroecology, metabolic rate, metabolic theory of ecology, mixed models, taxonomy, variance, WBE model.

INTRODUCTION

The metabolic theory of ecology (MTE) provides a powerful framework for explaining properties of whole organisms in terms of physiological processes (West et al. 1997; Gillooly et al. 2001; Brown et al. 2004). Many life history parameters are predicted to scale allometrically with body size by a power law relationship in which the exponent is a multiple of 1/4. MTE, however, remains highly controversial (Martinez del Rio 2008). Proponents claim MTE has wide support and can provide conceptual unification in ecology (Brown et al. 2004). Detractors have criticized the assumptions and biological relevance of MTE, as well as the supporting evidence and its claim to be mechanistic (Kozlowski & Konarzewski 2004; O’Connor et al. 2007).

The allometric scaling exponent for metabolic rate has been a particular focus for this controversy (Dodds et al. 2001; White & Seymour 2003; Savage et al. 2004; Farrell-Gray & Gotelli 2005; Glazier 2005; Nagy 2005). This debate has focused on two separate but related issues: whether the metabolic scaling exponent takes a particular value (e.g. 3/4 or 2/3), and whether the variability among taxa is significant. At the heart of metabolic theory is a model proposed by West, Brown and Enquist (West et al. 1997, 2000, henceforth WBE), which is based on the biophysical properties of the vascular system and predicts 3/4 power scaling in metabolic rate. Proponents of MTE have presented evidence supporting this prediction (Gillooly et al. 2001; Savage et al. 2004) and claim that deviations from 3/4 are either predictable exceptions (Allen & Gillooly 2007) or have limited generality (e.g. small numbers of species, limited range in body size: Savage et al. 2004; Farrell-Gray & Gotelli 2005). Critics have presented numerous counter-examples and concluded that the observed heterogeneity is sufficient to reject the notion that one ‘universal’ exponent applies to all organisms (Kozlowski et al. 2003; Glazier 2005; White et al. 2006, 2007; Capellini et al. 2010). If scaling is not universal then either WBE is incomplete (Martinez del Rio 2008), or perhaps other theories of metabolic scaling should be preferred.

In fact, WBE does predict that metabolic scaling exponents should vary, albeit within a fairly narrow range. West et al. (1997) predicted a steeper scaling exponent among smaller-bodied organisms due to the higher proportion of the vascular system that is made up of cubic branching vessels (e.g. capillaries), compared with the
area-preserving branching of larger vessels. Savage et al. (2008) explored this in detail by relaxing the assumptions of WBE: they predicted that an increase in body size should be associated with either a curvilinear decrease in scaling exponent and/or a decrease in the variance in scaling exponents (but see Kozlowski & Konarzewski 2005).

A number of competing theories have been proposed to explain why metabolic scaling should vary. Banavar et al. (2002) proposed that variation derives from inefficiencies in the balance between supply and demand for metabolites, and from physiological compensating mechanisms. Two theories make explicit predictions about the kinds of organisms in which deviations should be found: the metabolic level boundaries (MLB) and the cell metabolism hypotheses. The MLB hypothesis states that the range of possible metabolic states is defined by physiological constraints (Glazier 2005, 2010). It predicts that taxa with high metabolic rates exhibit relatively low scaling exponents, which is related to an older suggestion that ectotherms display steeper scaling than endotherms (Phillipson 1981). The cell metabolism hypothesis (Kozlowski et al. 2003) posits that whole organism metabolic rate is simply the sum of the metabolic rates of constituent cells, and that the metabolic rate of individual cells is determined by their size. In this model, metabolic rate scaling is close to isometric among organisms of the same mean cell size, but much shallower (~2/3 power) where body size differences are entirely due to variation in cell size. This leads to the prediction of a negative relationship between metabolic scaling and genome size scaling, as genome size is an important determinant of mean cell size (but see also Starostova et al. 2009).

In this paper, we provide a systematic description of the nature and magnitude of the variation in metabolic scaling exponents among animals. We provide the first quantitative estimate of the magnitude of variation in metabolic scaling among taxa, and test which taxonomic level(s) exhibit most variation. We use these findings to test two predictions from WBE, namely that metabolic scaling is universal and converges on 3/4 power. We then test a suite of hypotheses proposed to explain the range of variation in scaling exponents: a generalized version of WBE (Savage et al. 2008), the MLB hypothesis (Glazier 2005, 2010) and the cell metabolism hypothesis (Kozlowski et al. 2003). Our primary aim in describing this variation is to understand the controversy surrounding metabolic scaling exponents.

**MATERIALS AND METHODS**

**Data**

We collated five datasets on resting metabolic rate (RMR) and body mass for mammals \((n = 626\) species in 22 orders: Savage et al. 2004), birds \((n = 336, 21: Bennett & Harvey 1987\) reptiles \((n = 74, 3: Bennett & Dawson 1976\), fish \((n = 69, 13: Clarke & Johnston 1999\) and soil invertebrates \((n = 137, 5: Meehan 2006\). Although field metabolic rate data are available for some of these groups, we used RMR for modelling because of the much larger availability of data and because RMR has previously been used to support the notion of a universal scaling exponent of 3/4 (Gillooly et al. 2001; Savage et al. 2004). All RMR data were converted to species mean values in joules per second using standard formulae (Gnaiger 1983). For reptiles, we used only data where RMR was measured at 20 °C. For fish, we used the ‘best’ estimate of species mean RMR provided by Clarke & Johnston (1999). For soil invertebrates, we standardized the estimates of RMR to 20 °C and for the mean mass of the species (following Gillooly et al. 2001). We did not control for body temperature in endotherms because many species in our dataset lack body temperature data (Clarke & Rothery 2008). The range in endotherm body temperature is just a few degrees and shows a strong phylogenetic pattern (Clarke et al. 2010), such that variation within families is extremely small (but see Appendix S1). All data were log-transformed prior to analysis and log [mass] was additionally centred on zero.

**Model selection**

We fitted a series of linear mixed-effect models in which the slope (scaling exponent) and intercept (normalization constant) varied among taxonomic groups (phylum, class, order, family) as random effects (see Appendix S1). We compared the fit of these models with ‘universal’ models that had a single slope but taxon-specific intercepts. This approach allows us to estimate simultaneously the mean scaling exponent and the variance among taxa (thus reconciling the two parallel debates about metabolic scaling exponents). The use of mixed models has two distinct advantages over other approaches to allometric slopes estimation (e.g. Warton et al. 2006). First, it allowed us to explore variation in scaling exponents without prior expectation of which taxonomic units to compare. Second, taxon-specific scaling exponents are estimated with ‘shrinkage’ (Pinheiro & Bates 2004), so the test for universal vs. heterogeneous scaling is not unduly influenced by outliers with small sample sizes (see Appendix S1 for further details).

Quantifying the variance in scaling exponents implies an over-arching null hypothesis that all taxa share the same universal scaling exponent. However, we have no *a priori* expectation of which taxonomic level (or levels) at which variation will be found. Indeed, the hierarchical data structure means that there is no unique universal model (see below). We therefore adopted a mixture of null-hypoth-
esis significance testing and information theoretic model comparison (Stephens et al. 2007). We employed a stepwise model-fitting procedure (see Appendix S1 for full details) that yielded a set of 27 candidate models, including the ordinary least squares regression, 15 universal models and 11 alternate (heterogeneous) models. We compared these models quantitatively using Akaike weights (Burnham & Anderson 2002), which are based on the AIC scores for all models in the candidate set. We report results from all models with weights > 0.0005, and estimate the overall mean scaling relationship using weighted model-averaging (Burnham & Anderson 2002). Parameters were estimated by restricted maximum likelihood (REML) using the lme4 package (Bates et al. 2008) version 0.999375-28 in R 2.8.1 (R Development Core Team 2008). Further analyses were based on the overall best-fitting model. We used a likelihood ratio test (LRT) to compare the best alternate model with the best universal model (these models are nested: Table 1) and estimated confidence intervals for the mean scaling exponent. We also tested whether our results are sensitive to the statistical methods used, to small variations among endotherms in body temperature, and to measurement error (see Appendix S1 for details).

Explaining variation in scaling exponents

We then tested predictions about the variation in scaling exponents among orders. In this section we describe these tests in increasing order of complexity. We were able to test most hypotheses with modifications to our best-fitting model and using LRT.

We tested the MLB hypothesis (Glazier 2005, 2010) by measuring the strength of correlation between order-specific slopes and intercepts. Our best model contains a term estimating the correlation between random effects for slope and intercept, so we tested the significance of this term using LRT by comparison with a model lacking this term (i.e. the correlation term is constrained to zero). However, the estimated correlation may be biased if the mass at which intercepts are estimated is far from the range over which the slopes are estimated (Glazier 2009, 2010). We therefore performed the test on a modified dataset in which log [mass] was centred on zero for each order (following Enders & Tofighi 2007), such that intercepts (i.e. metabolic level) are the fitted values of log [RMR] at the geometric mean mass of each order. This procedure for estimating the strength of correlation is more robust than comparing slopes and intercepts extracted from a mixed-effects model (Kliegl et al. 2009). We compared the scaling exponents of endotherms and ectotherms (Phillipson 1981) by adding fixed-effect terms for the mode of thermoregulation and its interaction with log [mass] to our best model. For this comparison, both models were fitted with maximum likelihood (ML), rather than REML (following Pinheiro & Bates 2004).

We fitted four additional models to test the prediction by Savage et al. (2008) that metabolic scaling is a decreasing function of mass. We first added a quadratic term (i.e.

Table 1 Mixed models of the metabolism allometry

<table>
<thead>
<tr>
<th>Rank</th>
<th>Slope effects</th>
<th>Intercept effects</th>
<th>beta</th>
<th>LogLik</th>
<th>AIC</th>
<th>∆AIC</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>O</td>
<td>P C O F</td>
<td>0.748</td>
<td>-521.56</td>
<td>1061.11</td>
<td>0</td>
<td>0.599</td>
</tr>
<tr>
<td>2</td>
<td>O</td>
<td>C O F</td>
<td>0.748</td>
<td>-523.51</td>
<td>1063.03</td>
<td>1.92</td>
<td>0.230</td>
</tr>
<tr>
<td>3</td>
<td>P</td>
<td>O P C O F</td>
<td>0.760</td>
<td>-521.47</td>
<td>1064.95</td>
<td>3.84</td>
<td>0.088</td>
</tr>
<tr>
<td>4</td>
<td>C</td>
<td>O P C O F</td>
<td>0.748</td>
<td>-521.54</td>
<td>1065.08</td>
<td>3.97</td>
<td>0.083</td>
</tr>
<tr>
<td>5</td>
<td>O</td>
<td>F P C O F</td>
<td>0.734</td>
<td>-542.51</td>
<td>1107.01</td>
<td>45.9</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>O</td>
<td>P C O F</td>
<td>0.734</td>
<td>-546.55</td>
<td>1107.10</td>
<td>46.0</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>F</td>
<td>P C O F</td>
<td>0.752</td>
<td>-552.42</td>
<td>1122.84</td>
<td>61.7</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>C</td>
<td>P C O F</td>
<td>0.765</td>
<td>-553.58</td>
<td>1125.17</td>
<td>64.1</td>
<td>0</td>
</tr>
<tr>
<td>9*</td>
<td>P</td>
<td>C O F</td>
<td>0.747</td>
<td>-562.83</td>
<td>1139.65</td>
<td>78.5</td>
<td>0</td>
</tr>
<tr>
<td>11</td>
<td>P</td>
<td>P C O F</td>
<td>0.755</td>
<td>-562.55</td>
<td>1143.09</td>
<td>82.0</td>
<td>0</td>
</tr>
<tr>
<td>12</td>
<td>O</td>
<td>P C O</td>
<td>0.741</td>
<td>-587.79</td>
<td>1191.58</td>
<td>130.5</td>
<td>0</td>
</tr>
<tr>
<td>13</td>
<td>O</td>
<td>P O F</td>
<td>0.753</td>
<td>-606.91</td>
<td>1229.81</td>
<td>168.7</td>
<td>0</td>
</tr>
<tr>
<td>27†</td>
<td></td>
<td></td>
<td>1.028</td>
<td>-2594.27</td>
<td>5194.55</td>
<td>4133</td>
<td>0</td>
</tr>
</tbody>
</table>

Rank denotes the rank fit among 27 candidate models (only 1 of 15 ‘universal’ models is shown). Slope and Intercept columns indicate the taxonomic levels at which each parameter was allowed to vary (P, phylum; C, class; O, order; F, family). Beta denotes the overall mean scaling exponent of metabolic rate on body mass, LogLik is the log-restricted-likelihood, AIC is the Akaike Information Criterion, ∆AIC is the difference in AIC compared with the best model, and Weight is the Akaike weight (an absolute measure of support that sums to 1 across all candidate models).

*Best-fitting ‘universal’ model.
†Ordinary least squares regression.
log [mass] to the fixed part of the model, and then to the random part. We then added a cubic term to the fixed part and finally to the random part. At each step, we used LRT to compare the fit of each new model with the previous one, using ML fits for pairs of models that differed in the fixed-effect part of the model, and using REML fits for models that differed in the random-effect part (following Pinheiro & Bates 2004).

For the other correlative tests, we estimated order-specific scaling exponents ($b_i$) from the best-fitting model (see Appendix S1 for details). To test whether scaling is more variable among small-bodied orders (Savage et al. 2008), we measured the correlation between mean[log [mass]] and $|b_i - 0.75|$. We tested the cell metabolism hypothesis by measuring the correlation between $b_i$ and order-specific scaling exponents for genome size (following Kozlowski et al. 2003), using data from the Animal Genome Size Database (Gregory 2008). We selected a model of the genome size allometry using the procedures described above, with the additional constraint that all candidate models contained a term for random slopes at the order level. We then extracted order-specific scaling exponents for genome size (see Appendix S1 for details) and estimated the correlation with $b_i$ using standardized (reduced) major axis (Warton et al. 2006).

RESULTS

We found strong evidence for heterogeneity in metabolic scaling. The best universal (single exponent) model is a poor description of the data when compared with models in which the slope varies among taxonomic groups (Table 1). Just 4/27 models have any statistical support (defined as an Akaike weight > 0.0005), all of which contain a random slopes term for orders. The best model overall has a different slope for each order but not other taxonomic levels. The evidence ratio (Burnham & Anderson 2002) favouring this model over any universal model is $> 10^{17}$ (LRT: $\chi^2 = 78.5$, d.f. = 2, $P < 0.0001$).

While our models strongly reject the notion that a single exponent can explain the metabolic scaling of all animals, our results do support a mean value very close to 3/4 predicted by West et al. (1997). The model-average slope is 0.749 and the best-fitting model has a slope of 0.748 (Table 1). Support for 3/4 power scaling is not dependent on allowing the exponent to vary among taxa: most universal models have a slope close to this value. The standard error around the estimate from the best model is 0.017, with 95% confidence intervals of 0.71–0.80. Order-specific exponents, however, exhibit a much larger range of values (Fig. 1): the estimated variance is 0.0111, indicating that 50% of orders are expected to have scaling exponents outside the range 0.68–0.82 and 5% outside the range 0.54–0.95.

The conclusion that scaling varies among orders is robust to several forms of uncertainty. Our findings are qualitatively unchanged (1) under more conservative assumptions about the statistical penalties for complex models (Table S1), (2) when analyses are restricted to ectotherms at 20 °C or the subset of endotherms that could be standardized to 39 °C (Table S2) and (3) measurement error in the estimates of metabolic rate (Fig. S1, see Appendix S1 for details).

We then tested several explanations for why orders vary in metabolic scaling, starting with predictions from the generalized version of WBE. We found no correlation between body size and variability in scaling exponents among orders ($r = 0.02$, $P = 0.89$). We also found no evidence for curvilinearity in metabolic scaling for the full dataset ($P > 0.2$ for all tests), but we did find a significant quadratic term in some classes. Both mammals ($\chi^2 = 6.44$, d.f. = 1, $P = 0.011$) and the Malacostraca ($\chi^2 = 8.47$, d.f. = 1, $P = 0.004$) show curvilinearity in the opposite direction to that predicted by WBE: metabolic scaling is steeper at larger body size. Only insects ($\chi^2 = 4.61$, d.f. = 1, $P = 0.032$) show nonlinear scaling in the direction predicted by WBE; a simple power law is sufficient to describe scaling in other classes. We found no evidence that orders with steeper scaling have lower metabolic rates (the MLB hypothesis: $\rho = 0.16$, $\chi^2 = 0.48$, d.f. = 1, $P = 0.49$, Fig. 2), nor when each of the eight classes are analysed separately ($P > 0.33$ in all cases). We also found no significant differences in the scaling of endotherms vs. ectotherms ($b = 0.06$, $\chi^2 = 4.28$, d.f. = 2, $P = 0.12$). Con-
trary to predictions, our test of the cell metabolism hypothesis revealed a weak positive relationship between metabolic scaling and genome size scaling ($b = 0.0094, 95\% \text{ CI: 0.0058, 0.013}$).

**DISCUSSION**

This study represents the most comprehensive quantitative analysis of scaling patterns in metabolism to date. Overall, we found that metabolic scaling converges almost exactly on a value of $3/4$ predicted by West *et al.* (1997) with confidence intervals excluding the Figure of $2/3$ expected from surface-area to volume considerations (Fig. 1). The results, however, clearly indicate that metabolic scaling is not universal but dependent on the taxonomic level at which it is measured. Although other studies have reached similar conclusions (Glazier 2005; Kozlowski & Konarzewski 2005; White *et al.* 2006, 2007; Capellini *et al.* 2010), our results provide the first estimate of the variance among taxa and the taxonomic level at which that variance is manifested.

What do our findings mean for metabolic theory and WBE? On the one hand, we have found strong support for $3/4$ scaling of metabolism on average, and there are relatively few orders for whom the $95\%$ prediction intervals exclude $3/4$ (Fig. 2). On the contrary, we showed that scaling is neither universal nor explicable by relaxing the assumptions of the underlying pseudo-fractal resource distribution model (West *et al.* 1997; Savage *et al.* 2008). Thus, we can say that WBE is incomplete at best. Most predictions of MTE do not assume a constant metabolic scaling exponent, but those which do should therefore be reassessed. But perhaps our concept of universal scaling exponents is a straw man. We have used universal scaling to mean ‘a single exponent for all’, but perhaps the term is analogous to another cornerstone of metabolic theory, life history invariants (Charnov 1997). Invariants are not fixed, but rather are uncorrelated with body size and the subset of biologically interesting invariants exhibit unimodal central tendency with a limited range of variation (Savage *et al.* 2006). Thus, ‘$3/4$ scaling of metabolism is universal’ should perhaps be read as shorthand for ‘$3/4$ scaling of metabolism is the central tendency’. This raises the question of how much heterogeneity is permitted before universality (or invariance) no longer holds. Semantic issues aside, our findings present a challenge to WBE: it is not sufficient to predict accurately the mean scaling exponent because other explanations for $3/4$ power scaling (as the central tendency) are at least as parsimonious (Banavar *et al.* 2002; Darveau *et al.* 2010).
al. 2002; Ginzburg & Damuth 2008). Facing this challenge has already begun: recent work has shown that relaxing the assumptions of WBE leads to a range of predictions about allometric relationships (Price et al. 2007; Savage et al. 2008). Price et al. (2009) showed that such flexible approaches outperform universal models in predicting plant morphological traits. We look forward to new derivations of WBE that explore multiple physiological constraints and make a wider range of testable predictions about scaling exponents (Martinez del Rio 2008).

We found no evidence to support two competing hypotheses about variation in metabolic scaling. We found no correlation between scaling exponent and metabolic level (the MLB hypothesis, Glazier 2005, 2010), in contrast with the negative correlations reported by Glazier and colleagues for diverse taxa (Glazier 2008, 2009, 2010; Killen et al. 2010). One explanation is that mass-specific metabolic rate (Glazier’s preferred measure of metabolic level) is a decreasing function of mass itself (because scaling is generally shallower than isometric). We have observed that scaling is steeper in larger mammals and mammals (see also Kozlowski & Konarzewski 2005; Clarke et al. 2010): thus, large-bodied taxa in these groups have both steeper scaling and lower mass-specific metabolic rates, leading to a negative correlation between mass-specific metabolic rate and scaling. However, this pattern is not widespread (indeed, insects show the opposite pattern), so additional explanations are required to explain the discrepancy in results. Much of the evidence for the MLB hypothesis comes from ontogenetic (i.e., intraspecific) scaling patterns (Glazier 2009, 2010; Killen et al. 2010), which we have not examined, and it is possible that different physiological processes operate at this level (Glazier 2005). Further exploration of the MLB hypothesis (both assumptions and predictions) is therefore keenly anticipated. However, we recommend that future tests of the MLB hypothesis define metabolic level, as here, using data centred within groups (Enders & Toftigh 2007), to avoid the problems associated with mass-specific metabolic rate. We also found no support for the cell metabolism hypothesis (Kozlowski et al. 2003), at least when genome size is applied as a proxy for cell size. In fact, we found a positive correlation between order-specific scaling of RMR and genome size, contrasting with the significant negative correlation reported by Kozlowski et al. (2003) among orders of birds and mammals (our results were unchanged using this subset). There are two broad reasons for this discrepancy. One is that heterogeneity among orders in genome size scaling is extremely small compared with the heterogeneity within (by 3 orders of magnitude, results not shown). The other is the particular set of orders used for comparison by Kozlowski et al.: the correlation in birds was driven by two outliers, both of which are based on extremely small samples (Charadrifor-

ues $n_c = 4$ species; Gruiformes $n_{MR} = 4$): removing either causes the relationship to become nonsignificant. In mammals, the correlation disappears if data from the Xenarthra are included ($n_{MR} = 15$, $n_c = 10$).

It is not clear why the variation in scaling should be found among orders and no other taxonomic level. Orders are somewhat arbitrarily-defined and differ greatly in age (Avise & Johns 1999), so we caution against over-interpreting this result. However, we can be reasonably confident that differences among orders reflect divergent responses to natural selection. An alternative explanation would be that imbalances in the supply and demand of metabolites can explain variation in scaling exponents (Banavar et al. 2002).

In this model, steeper scaling ($b > 0.75$) derives from inefficient network design, but it does not seem reasonable to suggest that entire orders are metabolically inefficient as each by definition is an adaptive radiation. By contrast, shallow scaling is said to derive from metabolic demand exceeding supply, which is thought to be most likely during ontogenetic development or among closely related species (Banavar et al. 2002).

We conclude that the controversy surrounding metabolic scaling exponents has arisen, in part, because researchers have tried to estimate the mean and variance in scaling exponents separately. One part of the literature has compared theoretical predictions ($2/3$ or $3/4$) with the (mean) scaling exponent across a large span of body masses (Dodds et al. 2001; White & Seymour 2003; Savage et al. 2004), without admitting the possibility that scaling might not be universal. Other studies have estimated the scaling exponent at finer taxonomic scales in order to demonstrate that certain groups are exceptions, or that scaling is heterogeneous (Glazier 2005; Kozlowski & Konarzewski 2005; Nagy 2005; White et al. 2006, 2007). Our study provides a synthesis of these parallel debates: the exponent in animals is $3/4$ on average, but the variation among taxa is extensive and real (i.e. departures from $3/4$ have strong statistical support and cannot be explained as exceptions). Estimating both mean and variance in scaling reveals that extremely steep ($b > 1$) or shallow ($b < 0.5$) scaling should not be unexpected, nor inconsistent with the overall ‘rule’ of $3/4$ power.

Our attempts to understand the variation in metabolic scaling have been greatly enhanced by testing multiple predictions using a single dataset, and we encourage this pluralistic approach to be adopted more widely (see also Price et al. 2009). Unfortunately, the search for a mechanistic understanding of metabolic scaling is hampered by the fact that candidate theories do not make predictions that are mutually exclusive. For example, convergence on $3/4$ power scaling could emerge from simple considerations about space-time dimensionality (Ginzburg & Damuth 2008), or as a mid-point between boundary conditions (Glazier 2005, 2010). Curve-fitting, however sophisticated, is unlikely to
provide resolution (McGill 2003). The lack of overwhelming support for any one theory indicates that considerable theoretical and empirical work remains to be carried out to explain the diversity of scaling patterns observed in nature, and the physiological and ecological factors that drive them.

ACKNOWLEDGEMENTS

We are grateful to Peter Bennett, Andrew Clarke, Peter Cotgreave, John Damuth, T. Ryan Gregory and Tim Meehan for making their data available, and to Douglas Bates, Mark Brewer, Stephen Freeman, Reinhold Kriegl and Andrew Robinson for statistical advice. We also thank David Atkinson, Tim Blackburn, Jim Brown, Isabella Capellini, Andrew Clarke, Brian Enquist, Kate Jones, David Storch and five anonymous referees for comments on previous versions of this manuscript. NI is supported by a NERC Fellowship (NE/D009448/2).

REFERENCES


**SUPPORTING INFORMATION**

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

**Figure S1** Simulation results.

**Table S1** Summary of the best model of metabolic scaling under four different information criteria and penalty terms.

**Table S2** Summary of the best model of metabolic scaling for ectotherms and endotherms, both with and without temperature-correction.

**Appendix S1** Methodological details and further analyses.

As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials are peer-reviewed and may be re-organized for online delivery, but are not copy-edited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.

Editor, David Storch
Manuscript received 8 January 2010
First decision made 3 February 2010
Manuscript accepted 19 February 2010

© 2010 Blackwell Publishing Ltd/CNRS