**ABSTRACT**

Life on earth spans a size range of around 21 orders of magnitude across species and can span a range of more than 6 orders of magnitude within species of animal. The effect of size on physiology is, therefore, enormous and is typically expressed by how physiological phenomena scale with mass. When $b \neq 1$ a trait does not vary in direct proportion to mass and is said to scale allometrically. The study of allometric scaling goes back to at least the time of Galileo Galilei, and published scaling relationships are now available for hundreds of traits. Here, the methods of scaling analysis are reviewed, using examples for a range of traits with an emphasis on those related to metabolism in animals. Where necessary, new relationships have been generated from published data using modern phylogenetically informed techniques. During recent decades one of the most controversial scaling relationships has been that between metabolic rate and body mass and a number of explanations have been proposed for the scaling of this trait. Examples of these mechanistic explanations for metabolic scaling are reviewed, and suggestions made for comparing between them. Finally, the conceptual links between metabolic scaling and ecological patterns are examined, emphasizing the distinction between (1) the hypothesis that size- and temperature-dependent variation among species and individuals in metabolic rate influences ecological processes at levels of organization from individuals to the biosphere and (2) mechanistic explanations for metabolic scaling that may explain the size- and temperature-dependence of this trait. © 2014 American Physiological Society. Compr Physiol 4:231-256, 2014.

**Introduction**

Living species span a size range of around 21 orders of magnitude, from the smallest single-celled micro-organisms (~0.1 pg) to giant sequoias *Sequoiadendron giganteum* and redwoods *Sequoia sempervirens* weighing several thousand tonnes. To put this size range in perspective, the earth weighs $6 \times 10^{27}$ g, which is about $10^{21}$-times heavier than an elephant. The difference in size between the largest and smallest organisms is therefore equivalent to the difference in size between the largest extant terrestrial vertebrate and the earth itself. Even relatively similar organisms span a large size range: the smallest adult vertebrate is the fish *Paedocypris proge netica*, which measures 7.9 mm long at maturity (216) and probably weighs less than 1 mg; the largest is the blue whale *Balaenoptera musculus* that weighs up to 190 t (294). During ontogenetic development, some animals can increase in size by at least six orders of magnitude (205, 280). Biological processes are tightly related to the physical dimensions of the system (cell and organism) they occur in, especially through the implications of surface area, volume and their interaction. The effect of size on the biology of organisms is, therefore, enormous. The study of how aspects of biology vary with size is called scaling; when a trait varies in proportion to body mass it is said to scale isometrically (from Greek, meaning “equal measure”). Surprisingly, however, many characteristics of animals do not vary in direct proportion to their size; such scaling is referred to as allometric (meaning “by another measure”).

That many aspects of biology vary allometrically with animal size was probably first formally recognized by Galileo Galilei (1637, cited by 356) and the study of allometric scaling has attracted considerable attention ever since (5, 44, 52, 97, 145, 178, 188, 221, 314, 357, 381, 386, 387, 404). The effect of mass ($M$) on an aspect of biology ($Y$) is commonly described using a power function of the form:

$$Y = aM^b$$

where $a$ represents the value of $Y$ for an animal of unit mass and $b$ represents the scaling exponent. When $b = 1$ the relationship is isometric and $Y$ is proportional to $M$; when $b = 0$, $Y$ is independent of $M$; when $b$ takes other values the relationship is allometric. For example, the volume ($V$) of an object with characteristic length $l$ is proportional to $l^3$ and its surface area ($SA$) is proportional to $l^2$. Thus, since $l \propto SA^{1/2} \propto V^{1/3}$ and $V \propto M$, it follows that surface area is proportional to mass$^{2/3}$. This is true for simple objects such as spheres and squares, but
also for organisms with relatively complex shapes (e.g., mammals: 333). An advantage of the power function for describing scaling relationships is that the relationship between $Y$ and $M$ becomes linear when expressed on log-log axes:

$$\log(Y) = \log(a) + b \log(M)$$

The relationship between surface area and body mass, therefore, appears as a straight line with a slope of 2/3 (Fig. 1). Allometric scaling with an exponent between 0 and 1, as is the case for body surface area, indicates that relative to $M$, $Y$ is smaller in large animals (e.g., the SA : V ratio decreases with size). When $b$ is greater than 1, on the other hand, large animals have relatively large values of $Y$. For example, the scaling exponent for the size of weapons and ornaments is typically in the range 1.5 to 2.5 (209), indicating that large animals have relatively larger weapons for their body size than small animals.

Although the use of power functions to describe scaling patterns is widespread (44, 52, 138, 314, 357), it must be borne in mind that models that violate rules for dealing with dimensions are useful only as empirical descriptions, because the meaning of a model is lost if it contains transcendental functions of variables that are not dimensionless (212). The dimensionality problem inherent in the standard two-parameter power equation can be solved by measuring mass with reference to a characteristic mass $M_0$:

$$Y = Y_0(M/M_0)^b$$

However, the relationship now has three rather than two parameters, and there is no natural reference value $M_0$ for weights (17, 212). The use of the standard power function $Y = aM^b$ also introduces the problem that $a$ takes units of $YM^{-b}$, and is, therefore, a function of $b$ (455). These shortcomings do not devalue the standard two-parameter power equations as a useful empirical description of scaling patterns, providing that the units of mass are defined and are used in a consistent and clear fashion. For example, if the units of $M$ in Figure 1 were not defined then the scaling relationship between surface area (SA, in m$^2$) and $M$ (surface area = 0.092 M$^{0.67}$) would be nonsensical because the value of SA estimated from $M$ would differ if $M$ was measured in mg, g, or kg. Providing that any scaling relationship is treated as conditional upon measurement of $M$ in consistent units, no such problems arise.

### Describing scaling relationships

Statistical descriptions of the patterns of the relationship between $Y$ and $M$ are key to the development of an understanding of the scaling of biological traits. A distinction that is not often made in the physiological literature is that between the various forms and levels of allometric scaling (451). In the morphological literature, three types are recognized: ontogenetic, static, and evolutionary (see, e.g., 62, 313, 371). Ontogenetic scaling considers growth trajectories of the relationship between $Y$ and $M$ for a single individual, although cross-sectional comparisons of multiple individuals from the same species at different stages of ontogenetic development are more common than longitudinal comparisons in the physiological literature (e.g., 205, 280, 363, 382, 389). Static scaling considers the relationship between $Y$ and $M$ among individuals of the same developmental stage within a species, such as within an instar for insects (e.g., 342, 408). Evolutionary scaling considers the relationship between $Y$ and $M$ among individuals of different species, again at the same developmental stage. The distinction between these forms of scaling is particularly important when comparing among models that have been proposed to explain scaling relationships (see below), because some models have different explanations for different forms of scaling, whereas others do not (180, 202, 213, 249, 462).

The most common statistical description of the relationship between $Y$ and $M$, irrespective of the form or level of scaling being considered, is a linear relationship between log($Y$) and log($M$), usually calculated using ordinary least squares (OLS) regression. This approach makes three key assumptions about the data: (1) that the $X$-variable, $M$, is measured without error, (2) that linear regression of log-log data is an appropriate statistical model for describing the relationship between $Y$ and $M$, and (3) that the data are independent. Depending on the data being analyzed, one or more of these assumptions are likely to be violated.

### Regression models—measurement error

Measurement error refers to any variation that causes an observed value to be randomly different from the “true” value.
of interest (192, 252). It is important to recognize that in this context measurement error includes not only the sources of variation that biologists usually regard as error (e.g., instrument and transcription error), but also encompasses genuine biological variation such as interindividual differences as well as dieil and seasonal variation (160, 307). Such interindividual differences present a problem for scaling studies, because although they are important (i.e., they are repeatable, heritable, and have consequences for performance: 29, 50, 56, 291, 437, 442), they nevertheless contribute variance to the species mean which is typically used for such analyses. In scaling studies, instrument error is likely to be small (307, 314), but other sources of variation are likely to be non-negligible (e.g., 22, 75, 276, 318). When the X-variable in a scaling study is measured with error (e.g., when each data point represents the mean mass of multiple individuals, multiple measurements of the same individual, or measurements of mass taken from the literature for individuals other than those for which the Y-variable was measured), a commonly advocated approach is to used reduced major axis (RMA) regression rather than OLS regression (226, 383). The use of RMA regression is increasing, both because of its utility in estimating functional relationships (226, 360, 383), and because of an increased appreciation of the assumptions of OLS regression. However, RMA regression also makes assumptions about the error distribution of the data, and this assumption may also be violated. Specifically, RMA regression assumes that that error variance in Y is equal to that in X. McArdle (253) suggests that OLS is the better technique to use as long as the error variance in the independent variable is less than one third of that in the dependent variable, but few scaling studies assess these error variances. In the case of the interspecific relationship between metabolic rate (MR) and M, the coefficient of variation (CV = s.d. divided by mean) of MR and M are reasonably similar (374), but interspecific differences in MR between species of similar size that are associated with ecology and life-history are rather large and amount to sixfold to sevenfold for endotherms and ~25-fold for ectotherms (e.g., 431, 441, 444). For at least these data, Monte Carlo simulations suggest that OLS regression is appropriate (429). OLS regression is also considered appropriate when the purpose of a scaling relationship is to enable prediction of Y from measurements of X (383), and when X is thought to be affecting Y, rather than the reverse (380).

**Regression models—data transformation**

Related to the problem of measurement error is the problem of choosing an appropriate regression form. Linear OLS regression of log-log transformed data is the most commonly applied approach, but a number of recent studies have discussed potential problems associated with this approach (163, 183, 184, 298-305, 304). The use of log-log transformation potentially introduces bias when estimating data on the original untransformed scale, and also potentially introduces bias into estimates of the elevation and exponent of scaling relationships. A possible solution to these problems is to fit the power equation \( Y = aM^b \) directly, using nonlinear estimation without log-log transformation (e.g., 300, 304). Nonlinear regression methods can be implemented in many standard statistical packages, and generally result in models that fit values for large species better than power functions fitted by linear regression of log-log transformed data (184). A nonlinear fit to untransformed data applies a model with additive error on the original (untransformed) scale

\[
Y = aM^b + \varepsilon
\]

whereas linear regression of log-log transformed data applies a model with additive error to the transformed data

\[
\log(Y) = \log(a) + b \log(M) + \varepsilon
\]

which, when back-transformed to the original scale, results in a model with multiplicative error

\[
Y = aM^{b+1}
\]

The extent to which this is a problem depends upon the data in question. Since many biological phenomena are inherently multiplicative (60, 135, 203), linear regression of log-log transformed ontogenetic or interspecific data may be appropriate (112, 159, 292, 329, 429, 454, 456). In such cases, nonlinear models may poorly fit data for small values of M (429).

Of course, the foregoing discussion regarding the best way to fit a power equation \( Y = aM^b \) to empirical data hinges upon the appropriateness of this relationship for describing the relationship between Y and M. Early proponents of the use of such power equations were Snell (381), who used power functions to describe the relationship between brain size and body size, and Krogh (221), who suggested that the metabolic rate of endotherms was proportional to \( M^b \) rather than \( M \); power functions have remained in widespread use since then (5, 44, 52, 145, 178, 188, 314, 357). While many relationships are well described by a power function, it may not be appropriate for all types of data. Strict adherence to a power function may, therefore, hinder examination of the relationship between Y and M (379). Not all species and traits conform to linear relationships on log-log axes, and relationships may be multilinear or curved on log-log axes. Examples include, but are not limited to, aspects of sexual size dimorphism (105), morphology (e.g. 26, 177, 182, 197, 345, 375, 388), maximum speed during running or swimming (e.g., 61, 72, 127, 447), metabolic rate (e.g., 95, 138, 165, 210, 300, 309, 429), and population density (e.g., 376). For example, a critical assumption of a two-parameter power function estimated by linear regression of log-log transformed data is that the original data conform with a power function having an intercept of 0 in a plot with arithmetic coordinates (304). This assumption is rarely checked and is not always met (300, 305). An alternative is a three-parameter power function, which
accommodates a nonzero intercept through the addition of a mass-independent constant \(c\):

\[ Y = a M^b + c \]

This statistical model can be estimated by nonlinear regression of untransformed data, and assumes a component of variation in \(Y\) that is independent of body mass \(c\), and a component that scales allometrically in proportion to \(M^b\).

**Phylogenetic signal in interspecific scaling**

Most methods for line-fitting (including both OLS and RMA) assume that each data point in the analysis is independent of the others. This assumption is likely to be violated for comparative studies including data for multiple species, because closely related species are likely to be more similar that distantly related ones. The problem also potentially applies to intraspecific studies of related individuals because closely related individuals are more likely to be similar that distantly related ones, though this is rarely acknowledged except in quantitative genetic studies that partition phenotypic variation into genetic and other components (106, 243; see also 153 for a demonstration of the link between quantitative genetic and comparative analyses). Violation of the assumption of independence in interspecific analyses leads to inflated degrees of freedom, increased Type I error rates, overestimation of the strength of regression relationships, and a significant increase in the variance of the scaling exponent estimate (119, 125, 160, 295, 343). Blomberg et al. (34) have shown that for studies with 20 or more species, most traits show significant phylogenetic signal, defined as the tendency for related species to resemble each other. The traits considered included morphological, physiological, ecological, and behavioral ones, and significant phylogenetic signal was evident even for traits that are thought to be evolutionarily malleable or subject to relatively strong environmental effects, as well as for those traits subject to high levels of measurement error (34). This demonstrates unambiguously that many types of comparative data show phylogenetic non-independence, and that comparative studies (including those aimed at understanding the effect of body size) should be analyzed in a phylogenetic context. Indeed, incorporating phylogenetic information in allometric analyses is essential because, when data show phylogenetic nonindependence, the inclusion of phylogenetic information can significantly alter estimates of the scaling exponent (e.g., 69, 432, 443), although this is not always the case (e.g., 116, 431). A common measure of phylogenetic signal is \(\lambda\) (119, 306), which normally varies between 0 and 1 and quantifies the degree to which trait evolution deviates from Brownian motion \((\lambda = 1)\). \(\lambda\) is mathematically equivalent to phylogenetic heritability, defined as the proportion of variance in a character that is explained by the relationship among taxa as given by the phylogeny (153, 181, 242), thereby further consolidating the conceptual link between the estimation of phylogenetic signal in comparative analyses and heritability in quantitative genetic analyses, since heritability is the ratio of additive genetic to total phenotypic variance.

A range of methods are available for undertaking phylogenetically informed analyses. These include phylogenetic generalized least squares (128, 146, 225, 251), independent contrasts (110), phylogenetic eigenvector regression (92), phylogenetic autocorrelation (63, 64), and phylogenetic mixed models (153, 181, 242), though not all methods explicitly incorporate evolutionary models (118). Many of these can be implemented using freely available software packages. For example, independent contrasts (110) can be implemented in Mesquite (244) using the Phenotypic Diversity Analysis Package module (277) and phylogenetic generalized least squares can be implemented in \(R\) (330) using the Analysis of Phylogenetics and Evolution package (310), or in Matlab using the Regressionv2 program (225). The “caper” (Comparative Analyses of Phylogenetics and Evolution in \(R\)) package in \(R\) implements both PGLS and independent contrasts (297).

Recent work has provided methods for incorporating measurement error into comparative analysis (192), for disentangling phylogenetic and spatial effects (120, 223), for conducting phylogenetically informed logistic regression (191), and has applied Markov chain Monte Carlo algorithms to phylogenetic generalized linear mixed models, thereby extending the generality of phylogenetically informed analysis to a range of non-Gaussian data distributions (153).

Although some discussion remains regarding the need to account for phylogeny in comparative analyses (230, 269, 266, 265, 424-426), the value of this approach is now well appreciated (119, 125, 160, 225, 336, 343). The phylogenetically informed approach requires a phylogeny, and although large trees including thousands of species are becoming more common (e.g., 27, 28, 121, 195, 327, 328), appropriate phylogenies are nevertheless unavailable for some groups. It is clear, however, that it is better to include an incomplete tree than no tree at all (90, 91, 326), especially if the poorly known aspects of the tree (e.g., soft polytomies and unknown branch lengths) are appropriately accounted for (reviewed in 125). Errors in topology, for example, are most problematic when near the tips of the tree, but such errors tend to be conservative in that they generally act to conceal real relationships (394). The best approach for conducting an allometric analysis with an incomplete tree is, therefore, to include phylogenetic information, but reduce the impact of topological errors by representing uncertain relationships as polytomies and conducting statistical tests with appropriately reduced degrees of freedom (124, 126, 325). The extent of phylogenetic nonindependence in the data can then be tested for using metrics such as \(\lambda\) (119, 306) and the \(K\)-statistic (34). Such tests have revealed that phylogenetic nonindependence is common in most biological data (34), but is not ubiquitous (e.g., 237, 255, 449). The degree of phylogenetic nonindependence for a given data set can be explicitly incorporated in the analyses by modifying the covariance matrix in phylogenetic generalized least squares to accommodate the degree of nonindependence.
in the data (e.g., 53, 98, 431-433). Such an approach ensures reliable estimation of the elevation ($a$) and scaling exponent ($b$) of allometric relationships.

**Size-dependent covariates**

Many traits in addition to metabolic rate are correlated with body mass, and the incorporation of these into scaling relationships can influence the values of the estimated parameters. For example, body temperature is weakly associated with body size in some mammalian lineages, and inclusion of body temperature in the relationship between metabolic rate and body mass decreases the estimated scaling exponent of metabolic rate (443). Similarly, in accordance with Bergmann’s rule (24), body size generally increases with latitude in endotherms (e.g., 14, 15, 32, 275), and some, but not all, groups of ectotherm (3, 25, 66, 85). Among species of mammals, body mass may also be associated with diet and life-history traits, including litter size and maximum longevity (199, 267), and these may also have confounding associations with metabolic rate (e.g., 267, 286, 393, 429, 444). Though most studies of metabolic scaling fail to do so, it is, therefore, important to consider other biotic and abiotic variables that may be associated with both mass and metabolic rate when describing the scaling of metabolic rate. In doing so, however, it is important not to conflate correlation and causation; for example, an animal’s body temperature is dictated by its metabolic rate and thermal conductance, so body temperature is not a cause of variation in metabolic rate. Nonetheless, variation in body temperature can be used to explain variation in metabolic rate statistically, and can, therefore, be used to improve predictions of metabolic rate beyond those that can be made on the basis of body mass alone. As always, manipulative experiments should be used to establish causation; the inclusion of covariates in scaling relationships can establish only correlations.

**Data requirements**

When compiling data for scaling relationships, a commonly raised question concerns the quantity of data required to generate a robust relationship, both in terms of the number of data points included in the data set and the mass range of species being considered. Previous studies of intraspecific scaling of metabolic rate have included mass ranges that vary from less than 1.5-fold (e.g., 79) to over 3-million-fold (280). Data included in studies of the interspecific scaling of metabolic rate with broad taxonomic representation now include data spanning 16 or 20 orders of magnitude range in body mass (87, 250, 436), a range that comes close to encompassing all living organisms. For mammals, interspecific metabolic scaling relationships including 150 species spanning three to four orders of magnitude variation in body mass, because species from high-latitude and cold environments often have higher metabolic rates than those from warmer environments (e.g., 4, 11, 186, 194, 238, 385, 391, 392, 430, 431, 449). Among species of mammals, body mass may also be associated with diet and life-history traits, including litter size and maximum longevity (199, 267), and these may also have confounding associations with metabolic rate (e.g., 267, 286, 393, 429, 444). Though most studies of metabolic scaling fail to do so, it is, therefore, important to consider other biotic and abiotic variables that may be associated with both mass and metabolic rate when describing the scaling of metabolic rate. In doing so, however, it is important not to conflate correlation and causation; for example, an animal’s body temperature is dictated by its metabolic rate and thermal conductance, so body temperature is not a cause of variation in metabolic rate. Nonetheless, variation in body temperature can be used to explain variation in metabolic rate statistically, and can, therefore, be used to improve predictions of metabolic rate beyond those that can be made on the basis of body mass alone. As always, manipulative experiments should be used to establish causation; the inclusion of covariates in scaling relationships can establish only correlations.

**Data requirements**

When compiling data for scaling relationships, a commonly raised question concerns the quantity of data required to generate a robust relationship, both in terms of the number of data points included in the data set and the mass range of species being considered. Previous studies of intraspecific scaling of metabolic rate have included mass ranges that vary from less than 1.5-fold (e.g., 79) to over 3-million-fold (280). Data included in studies of the interspecific scaling of metabolic rate with broad taxonomic representation now include data spanning 16 or 20 orders of magnitude range in body mass (87, 250, 436), a range that comes close to encompassing all living organisms. For mammals, interspecific metabolic scaling relationships including 150 species spanning three to four orders of magnitude variation in body mass are sufficient to distinguish between scaling exponents of 2/3 and 3/4 (446), but less data may be sufficient for many purposes. Monte Carlo simulations suggest that variation in the calculated scaling exponent decreases most dramatically as the mass range approaches one to two orders of magnitude, suggesting that a data set spanning this mass range is sufficient to reliably estimate the value of $b$ (Fig. 2A; see also Ref. 76 who presented similar relationships for the population density of mammals). Similarly, the increase in precision associated with increases...
of sample size above approximately 100 suggests that this number is sufficient (Fig. 2B). It must be borne in mind, however, that while 100 species spanning approximately one to two orders of magnitude range in body mass may be sufficient to reliably estimate the scaling exponent for a given data set, extrapolation beyond the mass range of the data is inappropriate, and considerably more data may be required to detect subtleties in the data, such as curvature, and small data sets may not provide sufficient power to detect differences among subsets within the data.

A related problem for comparative studies concerns the number of individuals that are required to be measured for a species mean to be included in a data set. In their analysis of metabolic scaling in birds McKechnie and Wolf (256) considered only data that unambiguously met the criteria for basal metabolic rate (BMR, the metabolic rate of a postabsorptive inactive nonreproductive adult endotherm thermoregulating in a thermoneutral environment during its inactive phase: 262, 314) and had sample sizes with \( n \geq 3 \), and showed convincingly that for such analyses including data measured under poorly defined conditions biases the outcome of scaling analyses. However, McNab (264) demonstrated that the estimated BMR of a single individual is usually within 2% to 8% of the species’ mean when the individual is measured eight or more times. Since the difference between individual mean BMR and the mean of individual means is minimized when an individual is measured eight times (264), future work should attempt to obtain eight measurements per individual if a reasonable estimate of a species mean is required from measurements of a single individual (such data could be complimented by calculation of a phylogenetically informed prediction for the species of interest: see “Predicting traits from body mass,” below). Similarly, the inclusion of data for rare, elusive, or endangered species with small sample sizes in comparative studies is warranted to maximize phylogenetic or spatial representation. Again, single individuals should be measured multiple times (eight, ideally), though the inclusion of data for individuals measured a smaller number of times is unlikely to be problematic, because such data are unlikely to introduce bias. Such data are likely to increase the variance of the data set, however. If this is suspected to be a problem, funnel plots of mass-independent residuals against sample size or log-transformed sample size can be used to determine if data should be weighted by sample size; such approaches are common in meta-analyses (see, e.g., 100, 434, 435).

### Interpreting scaling relationships

**Interpreting the value of the scaling exponent**

Once a data set has been gathered and the scaling relationship determined, the interpretation of the relationship depends on the question at hand. For some studies, the value of the scaling exponent itself may be of interest, and may be compared to a null expectation to test a given hypothesis. For example, the role of gravity in the cardiovascular systems of terrestrial animals has been of interest for decades (16, 41, 129, 130, 174, 175, 231, 232, 278, 311, 362, 364-366, 448). Since the total height of the blood column above the heart increases with animal size, and the hydrostatic pressure at the bottom of a column of fluid is calculated as the product of fluid density (\( \rho \)), gravitational acceleration (\( g \)) and the height of the column (\( h \)), then blood pressure should be positively related to body size if the heart must work against gravity to pump blood to the head. If the heart does not work against gravity, then blood pressure should be size-invariant. The hypothesis that the heart works against gravity can therefore be tested by examining the scaling of systolic and diastolic blood pressure in terrestrial mammals, and testing for a scaling exponent that is significantly greater than zero (448). Such an analysis reveals that blood pressure increases with size in accordance with the predicted scaling of head-heart distance, in favor of the hypothesis (Fig. 3).

A similar approach is taken to test hypotheses concerning any hypothesized value of the scaling exponent. For example, because the weight of a bone varies in proportion to its volume (\( b = 1 \)), but its strength varies in proportion to its cross-sectional area (\( b = 0.67 \)), it can be hypothesized that the skeletal elements of large terrestrial species should be relatively more robust than those of small species to support their mass against gravity. Thus, it can be predicted that the skeletal mass of animals that must support their weight against gravity should increase allometrically (\( b > 1 \)), because the cross sectional area, not the mass, of bones must increase in proportion to animal mass. This hypothesis can be tested by examining the scaling of skeletal mass with body mass for terrestrial animals. When all available data for skeletal mass are considered; however, the scaling of skeletal mass for terrestrial animals ranging in size from 7 g arctic shrews *Sorex arcticus* to 9 t African elephants *Loxodonta africana* does...
 scaling relationships: Similar exponents for a single trait

One of the most common and valuable uses of scaling relationships lies in their ability to compare traits between groups of animals while accounting for the allometric relationship between the trait of interest and body mass. When the scaling exponents for two or more groups of animals are not significantly different, analysis of covariance (ANCOVA) can be used to test for differences between the groups to determine if there are mass-independent differences in the mean value of the trait of interest (302, 303, 383, 456). The major assumptions of ANCOVA are that (i) data are randomly selected from the population and randomly assigned to groups, (ii) within group regressions are homogenous, (iii) the covariate and treatment are statistically independent, (iv) covariate values are fixed and error free, (v) within-group regression are linear, (vi) conditional Y scores are normally distributed, (vii) conditional Y scores show homogeneity of variance, (viii) treatment levels are fixed (185). It is important to recognize that the ANCOVA approach is not equivalent to comparing mean values of mass-specific data (Y_{m-s} = Y/M), because Y_{m-s} is independent of M only when Y scales isometrically (i.e., when Y is proportional to M\(^{1}\)). When Y scales allometrically with M, the scaling exponent of Y_{m-s} (b_{m-s}) is the negative complement of the whole-body one (b_{body} = b - 1). When b \neq 1, the error introduced by using mass-specific data to account for variation in mass is substantial, even when the mass range is small. For a twofold range in M, for example, the error is 16 or 20% if b = 0.75 or 0.67, respectively.

The ANCOVA approach has been used in a phylogenetic context to compare the metabolic rates of arid and nonarid species on multiple occasions (e.g., 255, 289, 405, 429, 428). Species living in arid environments may be predicted to have low metabolic rates because resources are sparse and widely distributed (239, 259, 281), or, in the case of endotherms, because of constraints imposed by high temperatures on heat dissipation (272, 385, 431). Since the scaling exponent of field metabolic rate (FMR) is not significantly different between arid and non-arid birds (405), it is possible to test for differences in FMR between them while simultaneously accounting for the effect of mass using ANCOVA. Such an analysis reveals that species from arid environments have, on average for a given body mass, lower metabolic rates than those from

not conform to a simple two-parameter power function, but instead to a curved relationship on log-log axes (Fig. 4). This finding highlights the importance of considering nonpower-function allometric scaling, but also highlights the potential influence of high leverage values on the interpretation of scaling patterns. In this case, excluding data for the two species of elephant leads to a different interpretation: if only species up to and including 113 kg Gorilla are considered, then the scaling exponent of skeletal mass is not significantly different from one (Fig. 4). These analyses suggest that the scaling of skeletal mass is isometric for most terrestrial mammals, as is also the case for cetaceans (236, 334) and fish (447), but the largest terrestrial mammals have relatively heavy skeletons. It is noteworthy, however, that the relationship in Figure 4 includes no data for terrestrial mammals between 113 kg and 6 t, so this conclusion should be reevaluated as more data become available.

**Figure 4**  Scaling of skeletal mass (\(M_s\)) and relative skeletal mass with body mass (\(M\)) in mammals. Relative skeletal mass is calculated by dividing skeletal mass by body mass. Data were for 73 species and were compiled from published sources (21, 241, 319, 322, 335), and matched to a supertree of mammals (28). Data were analyzed using phylogenetic generalized least squares (PGLS) (128, 146, 251) in the APE (Analysis of Phylogenetics and Evolution) package (310) within R (189) according to established procedures (98, 156, 432). In addition to the dated branch lengths associated with the supertree, a range of branch length transformations were compared: star, log, punctuated, Grafen’s (146), Nee’s (324), and Pagel’s (308). For each of these models, a measure of phylogenetic correlation, \(\lambda\) (119, 306), was estimated by fitting PGLS models with different values of \(\lambda\) and finding the value that maximizes the log likelihood. The degree to which trait evolution deviates from Brownian motion (\(\lambda = 1\)) was accommodated by modifying the covariance matrix using the maximum likelihood value of \(\lambda\), which is a multiplier of the off-diagonal elements of the covariance matrix (i.e., those quantifying the degree of relatedness between species). All models were compared on the basis of Akaike’s Information Criterion (AIC) as a measure of model fit (49). The solid line is a significant second-order polynomial regression relating log(\(M_s\)) to log(\(M\)) and [log(\(M\)) + 4]\(^2\): the best model included a phylogeny with all branch lengths equal to 1 (\(w = 0.18, \lambda = 0.32\)); log(\(M_s\)) = −1.53 + 0.87 log(\(M\)) + 0.02 [log(\(M\)) + 4]\(^2\). Dashed line is the best model for data excluding Loxodonta africana and Elephas maximas and including a phylogeny with all branch lengths equal to 1 (\(w = 0.68, \lambda = 0.37\)), which is a linear model: log(\(M_s\)) = −1.19 + 1.02 log(\(M\)). The 95% confidence interval of the scaling exponent for the linear model includes 1 (95%CI: 0.98-1.06).
Comparing scaling relationships: Dissimilar exponents for a single trait

In the previous examples, mean values of a trait for two groups were compared while accounting for the relationship between the trait and body mass by including body mass as a covariate in ANCOVA. A requirement of ANCOVA is that the relationship with the covariate is uniform across groups; that is, the regression slopes are homogenous (assumption ii of ANCOVA, above). In practice, a prerequisite for ANCOVA is, therefore, demonstration that the scaling exponent of the trait is not significantly different between groups. When the scaling exponents differ between groups, it is not possible to compare them using ANCOVA (456). Such a problem arises when comparing the heart masses of mammals and birds (Fig. 6), the BMRs of wild and captive birds (255) and the FMRs of mammals and reptiles (289, 385). Demonstration of scaling exponents that differ significantly between groups can be regarded as evidence of a significant treatment effect (73, 301), leading to the conclusion that the groups are significantly different, but in some situations, it may be valuable to analyze the data further. In the case of heart mass, for example, Seymour and Blaylock (362) demonstrated significantly different scaling exponents between mammals and birds, and concluded that “the bird hearts were obviously heavier within the range of similar body mass” and that “the scaling factor was twice as high at a body mass of 1 kg, but the data converge in larger species” (p. 395). In such a situation it would be valuable to determine the range of body masses over which the heart masses of birds and mammals are different, and the range of masses over which they are not. Such an analysis can be undertaken by applying the Johnson-Neyman technique (196) demonstrates that at $P = 0.05$, regression elevations are not significantly different at masses above 4.26 kg (427), thus confirming the conclusion that the hearts of flightless birds are not significantly larger than those of similarly sized mammals (362). The vertical dashed line represents the lower limit of the region of non-significance. Within the region of nonsignificance there is no significant difference in elevation between the scaling relationships, so in this example the groups differ significantly in elevation to the left of the vertical line. Birds: heart mass ($g$) = 8.08 $M^{0.91}$, mammals: heart mass = 4.04 $M^{1.06}$.
the fields of medical and behavioral science, sociology, and ecology (e.g., 96, 132, 185, 227, 457), and more recently to the field of comparative physiology, including: analyses of growth and development in birds (347, 367) and marsupials (113, 395); metabolic physiology in birds (12, 157), mammals (428), and fish (176, 296); and digestive physiology in birds and mammals (59, 225). Application of the Johnson-Neyman technique to the heart mass data provided by Seymour and Blaylock (362) yields a region of nonsignificance that ranges from a lower mass of 4.26 kg to an upper mass well beyond the largest animal in the data set (Fig. 6). Thus, the analysis confirms the conclusion that the hearts of flightless birds are significantly larger than those of similarly sized mammals. The hearts of small (<4.26 kg) birds, on the other hand, are significantly larger than those of similarly sized mammals.

The major assumptions of the Johnson-Neyman technique are similar to those of ANCOVA (185): (i) the residuals of the within-group regressions of Y on X are independent, and individuals have been randomly selected from a specified population and randomly assigned to groups; (ii) the residuals are normally distributed; (iii) the residuals have homogeneous variance for each value of X; (iv) the residuals have homogeneous variance across treatment groups; (v) the regression of Y on X is linear; (vi) the levels of the covariate are fixed; and (vii) the covariate is measured without error. Assumption (i) is likely to be violated for comparative data because closely related species tend to be more similar than distantly related ones, but unfortunately there is not yet a phylogenetically informed implementation of the Johnson-Neyman technique. Nevertheless, the technique has been applied in studies that incorporate phylogenetic information wherever possible (225, 428). It is also unfortunate that the technique is not implemented in commercially available statistical packages. A Microsoft Excel spreadsheet that performs both ANCOVA and the Johnson-Neyman technique accompanies an earlier presentation of the technique (427) and is available for distribution via email from CRW. The limits of the region of nonsignificance are calculated according to

\[
X_{\text{lower}} = \frac{-B - \sqrt{B^2 - AC}}{A}
\]
\[
X_{\text{upper}} = \frac{-B + \sqrt{B^2 - AC}}{A}
\]

Where

\[
A = \frac{-F_{(a,1,N-4)}}{N - 4} (SSres_i) \left( \frac{1}{\sum X^2} + \frac{1}{\sum Y^2} \right) + (b_1 - b_2)^2
\]
\[
B = \frac{F_{(a,1,N-4)}}{N - 4} (SSres_i) \left( \frac{\bar{X}_1}{\sum X^2} + \frac{\bar{X}_2}{\sum X^2} \right) + (a_1 - a_2)(b_1 - b_2)
\]
\[
C = \frac{-F_{(a,1,N-4)}}{N - 4} (SSres_i) \left( \frac{N}{n_1n_2} + \frac{\bar{X}_1^2}{\sum X^2} + \frac{\bar{X}_2^2}{\sum X^2} \right) + (a_1 - a_2)^2
\]
\[
SSres_i = \left( \sum Y^2 - \frac{(\sum XY)^2}{\sum X^2} \right) + \left( \sum y^2 - \frac{(\sum xy)^2}{\sum x^2} \right)
\]

\[
F_{(a,1,N-4)} = \text{critical value of } F \text{ statistic at } \alpha \text{ for } 1 \text{ and } N - 4 \text{ degrees of freedom; } N = \text{total number of observations} = n_1 + n_2; \ n_1, n_2 = \text{number of observations in groups 1 and 2, respectively; } \bar{X}_1, \bar{X}_2 = \text{covariate means for groups 1 and 2, respectively; } a_1, a_2 = \text{regression intercepts for groups 1 and 2, respectively; } b_1, b_2 = \text{regression slopes for groups 1 and 2, respectively. The quantities } \sum x^2, \sum y^2, \sum xy, \sum x^2, \sum xy_1, \text{ and } \sum xy_2 \text{ are calculated according to:}
\]

\[
\sum x^2 = \sum X^2 - \frac{(\sum X_1)^2}{n_1}
\]
\[
\sum y^2 = \sum Y^2 - \frac{(\sum Y_1)^2}{n_1}
\]
\[
\sum xy = \sum XY - \frac{(\sum X_1)(\sum Y_1)}{n_1}
\]
\[
\sum x^2 = \sum X^2 - \frac{(\sum X_2)^2}{n_2}
\]
\[
\sum y^2 = \sum Y^2 - \frac{(\sum Y_2)^2}{n_2}
\]
\[
\sum xy = \sum XY_2 - \frac{(\sum X_2)(\sum Y_2)}{n_2}
\]

Comparing scaling relationships: Multiple traits and the analysis of residuals

In the previous examples, scaling exponents for a single trait (e.g., metabolic rate) were compared between groups of species (e.g., birds and mammals, or mammals from arid and nonarid environments) as a prelude to a comparison of these groups using ANCOVA or the Johnson-Neyman Technique. A further application of allometry lies in the comparison of scaling exponents for separate traits for a single group of animals to test hypotheses that the traits are related. Such an approach is often applied to test hypotheses of animal design, such as the hypothesis that heat loss through the body surface influences the scaling of metabolic heat production, as first proposed by Sarrus and Rameaux in the 1830s (39, 339, 346). Their logic was as follows: since the heat produced as a by-product of metabolism must ultimately be lost through the body surface, the rate at which heat is produced by animals (i.e., their metabolic rate) should be matched to the area over which it is dissipated. Thus, since body surface area scales as \(M^{0.67}\) (Fig. 1), the scaling exponent of metabolic rate should be similar. Indeed, a number of studies have reported scaling exponents close to 0.67 for endotherms (168, 170, 221, 339, 346, 443), though other studies have reported different values (23, 39, 40, 53, 206, 207, 210, 240, 350, 374, 432). It is important to recognize, however, that demonstration of statistically similar scaling exponents for multiple traits demonstrates only that they share a similar relationship with body mass; it does not demonstrate that the traits are functionally related (see, e.g., 271). Demonstration of a functional association requires demonstration that traits are related independent
of their shared relationships with body mass. In the case of the putative relationship between metabolic rate and body surface area, it must be demonstrated not that body surface area and metabolic rate scale with similar exponents, but that species with a relatively high body surface area also have a relatively high metabolic rate. Such associations are often examined using mass-independent residual values that are calculated by subtracting the trait value predicted by a scaling relationship from the measured trait value, and testing for an association between the traits (e.g., 444). Such an approach has been criticized as a statistically flawed ad hoc procedure (117, 123, 164). The most strident criticisms are that parameter estimates are biased and the error degrees of freedom in the analysis of residuals are overestimated because the estimation of the regression coefficients for the traits is not considered. An approach preferable to the analysis of residuals is to use standard or phylogenetically informed multiple regression to control for the potentially confounding effect of body mass (117). Application of this approach to test for an association between BMR and the body surface area of mammals demonstrates that these traits are not associated (444), despite scaling similarly with body mass (Table 1). It should be borne in mind, however, that just as size-dependent covariates can influence the relationship between traits and body mass (see “Size-dependent covariates,” above), mass-independent associations among traits may also be influenced by biotic and abiotic covariates (i.e., a mass-independent relationship between surface area and metabolic rate could be obscured by differences in ambient temperature among species and studies, for example).

The multiple regression approach has also been applied to tests of the symmorphosis hypothesis that animals are designed optimally (411, 414, 415). In this case, one might hypothesize that the maximum aerobic metabolic rate of an animal during exercise (\( \dot{V}O_2\text{max} \)), which represents the maximum rate of oxygen transport through the oxygen cascade from the lungs to the mitochondria, should be matched to capillary volume of the locomotory musculature (\( V_{cap} \)), the location of the final convective step of the oxygen cascade, and to total mitochondrial volume (\( V_{mt} \), where the oxygen is consumed). When the scaling of these traits is examined, they are found to scale similarly with body mass (\( b = 0.962, 0.984, \) and 0.956 for \( \dot{V}O_2\text{max}, V_{cap}, \) and \( V_{mt} \), respectively) (412). Again, however, this is not sufficient to demonstrate functional associations. Multiple regression indicates that \( \dot{V}O_2\text{max} \) is significantly correlated with both \( V_{cap} \) and \( V_{mt} \) independent of their shared relationship with body mass (Table 2), supporting the hypothesis that \( \dot{V}O_2\text{max} \) is associated with the aerobic capacity of the locomotory musculature (412).

The preceding examples concern cases where putative associations are examined between traits that scale similarly with body mass, but associations can also exist between traits that scale with significantly different scaling exponents. If the BMR of mammals is assumed to scale as a simple power function of mass, the scaling exponent is \( \sim 0.71 \) (53, 374, 432; see 71, 210, 283, 429 for discussion of curvature in the scaling of metabolic rate). Heart rate, on the other hand, scales with an exponent of \(-0.26 \) (Fig. 7). Heart rate and metabolic rate are related according to the Fick equation (111), which describes the relationship between rate of oxygen consumption (\( \dot{V}O_2 \)), heart rate (\( f_H \)), stroke volume (\( V_s \)), and the oxygen contents of arterial (\( C_{aO_2} \)) and mixed venous blood (\( C_{vO_2} \)):

\[
\dot{V}O_2 = f_H V_s (C_{aO_2} - C_{vO_2})
\]

According to the Fick equation, a high metabolic rate (high \( \dot{V}O_2 \)) is matched to an increase in either or both of heart rate and oxygen pulse (= \( V_s \) multiplied by oxygen extraction: \( C_{aO_2} - C_{vO_2} \)). The positive relationship between heart rate and metabolic rate underpins the “heart rate” method for estimating the energy expenditure of free-living animals (51, 148). Similarly, it can be hypothesized that species with high rates of metabolism and oxygen consumption will have high heart rates, but scaling relationships alone fail to support this hypothesis: metabolic rate increases with size (\( b = 0.71 \)) whereas heart rate decreases with size (\( b = -0.26 \)). Thus, the

---

**Table 1** Parameter estimates for the relationship between basal metabolic rate (BMR, mL h\(^{-1}\)), body mass (\( M, g \)), and body surface area (\( SA, m^2 \)) for mammals (\( \log BMR = 1.04 + 0.87 \log M - 0.30 \log SA \))

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>SE</th>
<th>( P )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1.04</td>
<td>0.27</td>
<td>0.001</td>
</tr>
<tr>
<td>( \log M )</td>
<td>0.87</td>
<td>0.21</td>
<td>0.0005</td>
</tr>
<tr>
<td>( \log SA )</td>
<td>-0.30</td>
<td>0.30</td>
<td>0.33</td>
</tr>
</tbody>
</table>

Data from Dawson and Hulbert (83), Reynolds (333), and White and Seymour (443).

---

**Table 2** Parameter estimates for the relationship between exercise-induced maximum aerobic metabolic rate (\( \dot{V}O_2\text{max}, mL min\(^{-1}\) \)) and body mass (\( M, g \)), muscle mitochondrial volume (\( V_{mt}, mL \)), and muscle capillary volume (\( V_{c}, mL \))

<table>
<thead>
<tr>
<th>Term</th>
<th>Estimate</th>
<th>SE</th>
<th>( P )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>0.56</td>
<td>0.26</td>
<td>0.06</td>
</tr>
<tr>
<td>( \log M )</td>
<td>-0.09</td>
<td>0.18</td>
<td>0.62</td>
</tr>
<tr>
<td>( \log V_{mt} )</td>
<td>1.10</td>
<td>0.18</td>
<td>0.0003</td>
</tr>
</tbody>
</table>

Data from Weibel et al. (412) and references therein.
largest animals have the highest absolute metabolic rates and the lowest heart rates, which seems counterintuitive. A test for an interspecific association between metabolic rate and heart rate remains valid; however, it is possible to test for an association between heart rate and metabolic rate, while accounting for body mass, using the multiple regression approach. Such a test reveals that heart rate and metabolic rate are indeed positively associated so, for a given size, species with high metabolic rates have high heart rates (Fig. 8). The discrepancy between the scaling exponents of heart rate ($b = -0.26$, Fig. 7) and metabolic rate ($b = 0.71$) can then be explained by the scaling of cardiac stroke volume ($b = 1.03: 362$), since the product of stroke volume and heart rate increases with size as $M^{0.77}$, which is similar to the scaling of metabolic rate.

While the multiple regression approach overcomes the problems associated with the analysis of residuals and has provided significant insight regarding the associations among metabolic rate and a range of other traits, it is nonetheless a potentially imperfect solution. This is because, as discussed above, many traits in addition to metabolic rate covary with body mass (52, 314, 357). This collinearity among traits can result in spurious conclusions about the relationship between dependent and independent variables in multiple regressions, because the partial regression coefficients associated with the independent variables may not be representative of the relationship that exists in the population (456). Such problems can be overcome by moving from a univariate approach where associations between a single dependent and multiple independent variables are examined to a multivariate approach in which covariances among traits are estimated in a mixed model framework that incorporates phylogenetic information (153).

Using scaling relationships

**Predicting traits from body mass**

When a trait of interest is significantly related to body mass, the regression describing the relationship can be used to predict the value of the trait, based on measurements of only body mass. This is an often quoted use of allometry (e.g., 149, 200, 235, 246, 247, 256, 267, 268, 287, 289) and is particularly prevalent in the human literature where scaling principles have been used to predict rates of metabolism (e.g., 172, 173, 284) and drug clearance (e.g., 396, 397, 398). Such analyses of intraspecific scaling for humans now routinely include data for metabolic rate and organ sizes of hundreds of individuals, and have been particularly useful for establishing the influence of body composition and stature on the scaling of metabolic rate with size (171, 172, 173, 284), and for estimating in vivo metabolic rates of organ-tissue compartments (409, 410).

Scaling relationships can also be applied to the prediction of physiological characteristics of extinct species (114, 270, 348, 359, 361, 368), and, because of the allometric relationship between body mass and other morphological variables, can be used to predict morphological characteristics that are
A scaling of basal metabolic for 148 species of murid rodent. Data adapted, with permission, from White and Seymour (443) synonymized to match the supertree of Bininda-Emonds et al. (28); see White et al. (432) for details. Solid blue line is the ordinary least squares relationship; dashed blue lines are the 95% prediction interval of the ordinary least squares (OLS) relationship. Solid red line is the relationship for Notomys alexis estimated by independent contrasts [PIC] (110) following Garland and Ives (128); dotted red lines are the 95% prediction interval for the phylogenetically informed regression. (B) BMR measured for 11 individual Notomys alexis (mean mass 33 g) using indirect calorimetry (440) shown ± SEM and compared with predicted BMR for a 33 g murid rodent (shown ± SEE) for the OLS and PIC regressions presented in (A). Error bounds of BMR value predicted by OLS encompass the measured value, but absolute BMR is overestimated by 29%, and the OLS relationship estimates BMR with considerable uncertainty. The PIC estimate of BMR for Notomys alexis is more accurate and overestimates BMR by only 4%, an error similar to the measurement error associated with experimental determination of metabolic rate by indirect calorimetry (e.g., 439), but the error bounds associated with the PIC estimate of BMR are wider than those of OLS. Note that the error bars for predicted BMR are asymmetric because of back-transformation from log-transformed data.

The standard error of estimate provides an overall indication of the accuracy with which the fitted regression function predicts the dependence of $Y$ on $X$ (456), and has desirable properties not shared by the coefficient of determination: for a given magnitude of residual variation, $r^2$ increases with the mass range of the sample, whereas $s_{YX}$ does not (Fig. 10). Quantification of residual variation in terms of standard error of estimate ($s_{YX}$) and residual mean square ($s^2_{YX}$) also permits calculation of the error associated with values of $Y$ predicted by $X$. The standard error of a predicted value of $Y$ for a given value of $X$ is (383):

$$s_Y = \sqrt{s^2_{YX} \left[ 1 + \frac{1}{n} + \frac{(X_i - \bar{X})^2}{\sum x^2} \right]}$$

in endotherms (56, 444), and over an order of magnitude in ectotherms (250, 441).

Smith (379) and McNab (261) suggest that the standard error of estimate ($s_{YX}$) provides a better measure of residual variation than the coefficient of nondetermination ($1 - r^2$), which expresses the proportion of variance of a variable that is not explained by another variable (383). $s_{YX}$ is calculated from the residual mean square $s^2_{YX}$:

$$s_{YX} = \frac{\sum (Y_i - \hat{Y_i})^2}{n - 2}$$

difficult or impossible to measure, such as the body mass of sharks (279) and extinct species (10, 136, 358). Making such predictions based on scaling relationships seems reasonable, because body mass typically accounts for most of the interspecific variation in a range of traits (e.g., metabolic rate: 429, 441, 445). It is important to recognize, however, that the coefficient of determination $r^2$ can be a poor descriptor of the goodness of fit of scaling relationships (76, 261, 314, 378); for a given magnitude of residual variation, $r^2$ increases with both the slope of a relationship and the $X$-range of the data, and high values of $r^2$ can conceal surprisingly large amounts of residual variation. For example, for the murid rodents in Figure 9A, BMR varies 84-fold from 13.5 mL h$^{-1}$ in least gerbils Gerbillus pusillus weighing 12.6 g (47) to 1141 mL h$^{-1}$ in Gambian pouched rats Cricetomys gambianus that weigh 1.9 kg (208). Among the 148 species of Muridae for which data are available, the coefficient of determination $r^2 = 0.79$, so mass alone explains 79% of the variation in BMR. Thus, just over one-fifth of the variation in BMR remains unexplained. This seems a relatively small amount, but is actually considerable. For example, fat mice Steatomys pratensis and silver mountain voles Alticola argentatus both have a mean body mass of 38 g, yet their BMRs differ by over sixfold (103, 416). When larger mass ranges are considered, it is not uncommon for mass to explain over 95% of the variation in metabolic rate, yet large residual differences between species are observed at all body masses and amount to several fold.
Comparing newly measured to existing data

When comparing newly obtained data for a species to those available for other species, a common approach is to compare the measured value to that predicted from a scaling relationship. The percentage of the species’ value relative to the “expected” value from the allometric equation for an animal of equivalent mass is then often interpreted to be relatively high (somewhat greater than 100%), relatively low (somewhat less than 100%), or approximately as expected (around 100%) (74). Such qualitative comparisons are potentially valuable, but are necessarily subjective because it is not clear how high or low new data need to be before the difference is important. For example, it has previously been hypothesized that mammals of >200g body mass that feed extensively on termites have metabolic rates lower than predicted from body mass (260, 263), presumably because a termite diet has a low net energy yield (1, 260, 322). Indeed, comparison of the BMRs of termitivorous numbats *Myrmecobius fasciatus* and aardwolves *Proteles cristatus* to the BMR predicted on the basis of their body mass reveals that numbats and aardwolves have BMRs that are 83.6% and 74.2% of that predicted by their respective body masses on the basis of scaling relationships for dasyurid marsupials and Carnivora, respectively (74). Similarly, the FMR of short-beaked echidnas *Tachyglossus aculeatus* varies from ~25% to 50% of that expected on the basis of body mass (38, 147, 288, 354). To determine if such differences are statistically significant; however, it is necessary to establish the 95% prediction confidence limits (74, 383, 456).

The 95% prediction confidence limits of a predicted value of *Y* (*Ŷ*) are calculated from the standard error of a predicted value of *Y* (*SŶ*) as:

\[
Ŷ \pm t_{(2, (n−2))}SŶ
\]

where *t* is the critical *t* value and *α* is the probability value (e.g., 0.05) for a two-tail *t* test for *n*−2 degrees of freedom (see also 124, 128 for phylogenetically informed implementations of this approach). If the newly measured value falls outside of the 95% prediction interval, then it can be considered to be significantly different from the remaining data at the specified *α* level. Prediction intervals tend to be wide (e.g., Fig. 9A), and are much wider than the 95% confidence interval of the regression mean. Thus, the BMRs of numbats and aardwolves are not significantly lower than those of other mammals (74), and a murid rodent weighing 33g would need to have a BMR lower than 52% of that predicted by mass to be significant different from the remaining murids (Fig. 9A). The power of such comparisons can be increased when data for multiple species are available (e.g., a comparison of multiple ant and termite eating species with other mammals), and groups of species are compared (see “Comparing scaling relationships,” above).
Metabolic Scaling

Perhaps the most widely examined physiological scaling relationship, and certainly the most controversial, is that between metabolic rate and body mass. Measures of metabolic rate integrate a wide variety of functions performed by animals (42, 155, 213, 390). Animals expend energy for many processes including the maintenance of homeostasis, foraging for and digesting their food, to cover overhead costs of processes such as growth, and to search for mates and reproduce. All of these aforementioned processes involve energy being used to do metabolic work (rather than being stored in the body as new tissue including growth or reproduction) and, therefore, contribute to the metabolic rate. Variation in metabolic rate is therefore linked to Darwinian fitness (see, e.g., reviews: 29, 50, 55, 211, 437). In larval radiated shanny Ulvaria subbifurcata and juvenile garden snails Helix aspersa, for example, standard metabolic rate (SMR, the metabolic rate of an inactive nonreproductive postabsorptive ectotherm measured at a known temperature during its inactive phase: 115, 314) is under a combination of directional and stabilizing selection such that individuals with low and intermediate metabolic rates are favored over those with high metabolic rates (13, 35). Low BMR improves starvation resistance in rats (338), but individual cockroaches Nauphoeta cinerea with low SMR do not live longer than those with high BMR under conditions of food and water restriction (351). Individual male Leach’s storm-petrel Oceanodroma leucorhoa with low BMR breed earlier and produce offspring that grow faster than individuals with high BMR (33), and individual female cockroaches with low SMR have shorter gestation durations than those with high BMR (352). BMR is also associated with over-winter survival in some, but not all, species (36, 37, 193, 224), and interindividual differences in metabolic rate are associated with differences in behavior (29). Metabolic rate has also been hypothesized to underlie a range of ecological patterns (9, 42, 46, 215, 273, 274, 285).

Metabolic rate is usually measured in a laboratory setting by direct calorimetry as heat production (201) or by indirect calorimetry as oxygen consumption or carbon dioxide production (228, 229, 417, 453). For free-living animals, metabolic rate is typically measured using either doubly labeled water (51, 370, 384) or heart rate (51, 148). Measurements of metabolic rate are now available for several thousand species (250, 385, 412), and studies of the scaling of metabolic rate have a history going back almost 200 years. The earliest work (Sarrus and Rameaux, 1838, cited in 39) suggested that because the heat produced as a by-product of metabolism must ultimately be lost through the body surface, the scaling exponent of metabolic rate should be similar to that of body surface area ($b = 0.67$, Fig. 1). This hypothesis was later supported by Rubner’s (346) measurements of heat production and body surface area in dogs, and a number of subsequent studies reported a scaling exponent close to 0.67 for the BMR of endotherms (168, 170, 221, 256, 315, 443).

Beginning with Kleiber (206, 207) and Brody and Proc- tor (40), however, other studies reported that metabolic rate scaled with an exponent greater than 0.67 in a wide array of animals (e.g., 23, 39, 167, 350), and by the 1980s, there was a broad consensus that metabolic rate scaled, on average, with an exponent of 0.75 (52, 314, 357). The perceived ubiquity of quarter-power scaling in turn led to the search for explanations for the origin of this non-Euclidean scaling (e.g., 214, 249, 421, 419), with the potential that explanations for metabolic scaling might also underlie scaling patterns observed in biochemical, physiological, and ecological systems (e.g., 42, 418, 422). The most recent phylogenetically informed studies of endotherms, however, have rejected both 0.67 and 0.75 as universal scaling exponents for these animals, and have stressed that no single scaling exponent adequately describes the mass dependence of metabolic rate (53, 255, 374, 432). Other studies have demonstrated that when plotted on log-log axes, the relationship between mass and metabolic rate for mammals is actually curved, such that the scaling exponent of metabolic rate increases with size (71, 95, 138, 165, 210, 309, 429).

Variation in the scaling exponent of metabolic rate is not limited to a single group of organisms, or to the resting state. The scaling exponent of metabolic rate differs between endothermic and ectothermic animals (69, 107, 315, 434, 441), and the scaling exponent for prokaryotes is higher than for other organisms (87). The scaling exponent of metabolic rate is also typically higher during exercise than during rest (140, 139, 141, 143, 205, 341, 412, 434, 449), and varies with habitat, phylogenetic affinity, and developmental mode in mammals (240, 283, 432, 444), with captivity in birds (255), with laboratory acclimation in scorpions (402), with lifestyle and temperature in fish (204), and with a range of other factors (138, 143, 437). A consequence of the variation in the scaling exponents for $\dot{V}O_2\text{max}$, FMR, and BMR (Fig. 11) is that aerobic scope ($= \dot{V}O_2\text{max} \text{ divided by BMR}$) tends to increase with size (30, 48), whereas activity scope ($= \text{FMR divided by BMR}$) tends to decrease (423). Thus, the largest animals have the greatest capacity to increase metabolic rate above basal levels, but free-living large animals routinely operate at metabolic levels only a few fold higher than basal (Fig. 12). Humphries and Careau (187) used scaling relationships to predict the influences of allometric scaling, air temperature, and mode of activity on the extent to which heat generated as a by-product of activity could be substituted for heat generated for thermoregulation. They predicted that the opportunity for heat substitution should increase with body size, leading to more similar metabolic rates between active and resting rate in large than small animals and thereby perhaps explaining the triangular pattern observed in Figure 12.

Mechanistic theories for metabolic scaling

The methods for scaling analyses presented thus far have focused largely on describing patterns of variation between physiological traits and body mass. Such exploratory analyses are valuable, because they document the association between
mass and the traits of interest and provide predictive equations that can be used to estimate the value of traits based on measurements of only body mass. Such descriptions do not, however, tell us why body mass is associated with traits in the manner that it is. For the case of metabolic rate, a wide range of mechanistic hypotheses have been proposed, beginning with the heat-loss hypothesis proposed by Sarrus and Rameaux that predicts metabolic rate should be proportional to body mass $^{2/3}$ (39, 339, 346). Many of these hypotheses are reviewed and critiqued in detail elsewhere (5, 6, 95, 138, 143, 314), and the debate concerning the factors that cause metabolic rate to scale allometrically with body mass remains one of the most enduring in biology (5, 6, 18-20, 45, 69, 82, 88, 89, 95, 99, 108, 109, 133, 137, 138, 143, 150-152, 168-170, 207, 214, 217-221, 257, 258, 312, 337, 340, 339, 369, 412, 413, 421, 419, 434, 449). Rather than repeat detailed critiques of these theories here, we instead highlight a small number of mechanistic theories that make quantitative predictions about the scaling of metabolic rate and use these to illustrate the difficulties involved in selecting between competing theories. We then discuss methods for evaluating the theories, and demonstrate that these theories can be used to provide a mechanistic basis for ecological and other patterns that are linked to metabolic rate.
Distribution network geometry

Probably the most well-known hypothesis proposed to explain the allometric scaling of metabolic rate is the theory of West and co-workers (418, 419, 422) that predicts the allometric scaling of metabolic rate based on the minimization of transport costs through fractally branching networks that distribute oxygen and nutrients. Although the classic prediction for metabolic scaling based on these models is a scaling exponent of $3/4$ (e.g., 419), the authors of the theories acknowledge that variation exists (e.g., 87) and the theory can predict a range of other values of the scaling exponent depending on the geometry of the networks (20, 210, 323, 349). Nonetheless, the canonical prediction of $3/4$-power scaling of metabolic rate continues to be used (e.g., 179, 462). An alternative model of nutrient transport predicts an exponent of 2/3 for the scaling of whole animal metabolic rate (94).

Dynamic energy budget theory

The Dynamic Energy Budget (DEB) theory (213, 214) considers metabolic rate to comprise a weighted sum of three processes: assimilation, growth, and dissipation. The latter category encompasses somatic maintenance (including foraging and locomotion), maturation, maturity maintenance and any overhead costs of reproduction. The theory is based on generalized surface area (source) and volume (sink) relationships, with body mass decomposed into two indirectly measurable state variables, the “reserve” and the “structure.” The composition of reserve and structure are assumed to remain constant (the “strong homeostasis” assumption) but may differ from each other. This compositional assumption enables a complete elemental analysis of the mass budget in terms of “macro-chemical equations” and provides a mechanistic underpinning to indirect calorimetry. Under constant food, the ratio of reserve to structure remains constant and hence so does the entire body composition (the “weak homeostasis” assumption). Energy and matter are assimilated in proportion to structural surface area (e.g., of cells or the gastrointestinal tract), and directed first to the reserve pool of the organism before being mobilized and allocated in fixed proportions to (i) the growth and maintenance of structure and (ii) maturation and maintenance of maturity, and to reproduction once puberty is reached.

A important distinction between DEB and other theories, and one that can be used to design empirical tests among DEB and other theories (202), is that the mechanisms invoked to explain intraspecific scaling relationships are different to those that explain interspecific scaling. Interspecifically, allometric scaling arises because the contribution of nonrespiring reserves to body mass increases with body size. For example, body fat (which is not strictly equivalent to reserve) scales as $M^{1.19}$ in mammals (52, 316) and has a very low mass-specific metabolic rate (102). Intraspecific metabolic rate, on the other hand, varies with size as a consequence of variation in the relative contributions of assimilation, maintenance, growth, and maturation during development. DEB theory predicts that metabolic rate scales with exponents between 0.5 and 1 for two-dimensional organisms, and between 2/3 and 1 for three-dimensional organisms, and predicts many other intra- and interspecific scaling relationships in addition to metabolic rate (e.g., 54, 213, 214, 293, 407). Notably, in the limit of infinite mass, DEB theory predicts $3/4$ power scaling, as does the more recent distribution network theory of West and co-workers, but for completely different reasons (214, 249).

Metabolic level boundaries

The metabolic level boundaries (MLB) hypothesis predicts variation in the scaling exponent of metabolic rate based on variation in the relative importance of two boundary constraints between groups and activity levels (138, 143). The boundary constraints are surface-area related effects on fluxes of metabolic resources, wastes, and (or) heat; and volume-related effects on energy use and power production. Volume-related effects scale isometrically with mass and apply at low and high levels of metabolic intensity, whereas surface-area-related constraints scale as mass$^{0.25}$ and are more prominent at intermediate levels of metabolism. Thus the scaling exponent of metabolic rate is predicted to vary in a U- or V-shaped pattern with metabolic intensity, bounded by 1 at low and high levels of metabolism and 2/3 at intermediate levels, in good agreement with data for a wide range of organisms (140, 139-143).

Heat dissipation limits

The heat dissipation limits (HDL) hypothesis predicts that the daily energy expenditure of endotherms is constrained not by their ability to acquire and process energy but by their capacity to dissipate the heat produced as a by-product of metabolic processes (385). The model builds on an elegant experiment demonstrating that lactating mice increase food intake and milk production when shaved (222), although the importance of heat stress for milk production had been demonstrated previously for large (>50 kg) domestic animals with low surface:area volume ratios (e.g., 31, 122, 385). Expansion of these experimental tests of the HDL theory beyond lactation to include a greater emphasis on exercise would be valuable (e.g., 459, 460); many studies have examined the influence of exercise on heat dissipation (e.g., 198, 233, 248, 321, 353, 399-401), and quantitative examination of this wealth of information may provide a useful basis opportunity for further tests of the HLD hypothesis. The HDL hypothesis predicts the scaling of metabolic rate for free-living endotherms on the basis of size-dependent variation in their maximal capacity to dissipate heat, yielding metabolic scaling exponents ranging from 0.47 to 0.50 (385), which is similar the scaling exponent of 0.46 predicted for the metabolic rate of thermoregulating furred endotherms at an ambient temperature 20°C (317).
Muscle aerobic capacity

During exercise-induced maximal aerobic metabolism most (>90%) metabolic activity is associated with work done by the locomotor muscles and delivery of substrates and oxygen to these (412). Given that most of the oxygen consumed during intense activity is used by the mitochondria of the locomotory musculature, it is reasonable to expect that $\dot{V}O_2\text{max}$ should scale in proportion to the total volume of mitochondria in the working muscles, and therefore, that the scaling exponent of $\dot{V}O_2\text{max}$ should approximate that of the scaling exponent of the total volume of mitochondria in the locomotory muscles. Weibel and colleagues (412) compiled estimates of total muscle mitochondrial volume for 11 species varying in size from 20 g to 475 kg, and report a scaling exponent of 0.956, which is very close to the scaling exponent of $\dot{V}O_2\text{max}$ for the same species (0.962).

Comparing theories for metabolic scaling

How should one select among the set of theories for the allometric scaling of metabolic rate described above? Each clearly provides a good fit to at least some of the metabolic data that they seek to describe, and so by the criteria of fit to data alone each must be considered acceptable, at least for some situations. However, comparing models on the basis of how well they predict available data will fail to distinguish between competing theories that predict the same value of the scaling exponent. For example, the original heat loss hypothesis for the scaling of metabolic rate predicts 2/3-power scaling of the BMR of endotherms; the same prediction is made by an alternative theory based on heat loss (339), a theory based on dimensional analysis and biological similarly (150, 151), two theories based on nutrient supply networks (20, 94), and the MLB hypothesis (138, 143). Similarly, the canonical 3/4-power scaling of metabolic rate can be predicted by the same theory of dimensional analysis and biological similarity, if a small empirical adjustment is made (150-152), as well by other theories based on biological similarity (99), elastic similarity (109, 257), nutrient supply networks (18, 19, 421, 419), and reserve/structure geometry (214, 249), among others. Given that the predictions of theories for metabolic scaling often overlap, it is therefore essential that the fit to data is not the only basis on which the theories are compared. Theories must also be evaluated on the legitimacy of the assumptions that underpin them (202).

A potential additional criterion by which competing explanations for metabolic scaling can be compared is their relative complexity, so that the best of two models that describe a given dataset equally well is the one that describes the data with the smallest number of parameters that must be estimated from the data. Indeed, simple explanations that incorporate a minimum of detail are sometimes regarded as more parsimonious than more complicated ones (461). Such ideas form the basis of several methods of formal model comparison (see 49), and applications of such approaches to the scaling of metabolic rate generally reveal that more complicated models describe the available data better than simple ones (e.g., 190, 436). Again, however, theories that predict the same scaling relationship with the same number of free parameters will be indistinguishable.

Given the problem of distinguishing between competing theories that make overlapping predictions, it is essential that the presentation of theories include clear descriptions of the unique predictions made by the theory, to facilitate tests that distinguish between alternatives (77, 162, 202, 372, 438). Such predictions should emerge from the theory, but will ideally complement the support provided to the theory by observed patterns of metabolic scaling, and will also ideally incorporate some form of experimental manipulation (138). For example, the HDL hypothesis has been tested by manipulating heat loss through fur removal or cold exposure (e.g., 222, 377, 458), the hypothesis that muscle aerobic capacity dictates the scaling of $\dot{V}O_2\text{max}$ could be tested by measuring the size dependence of training-induced changes in muscle mitochondria and quantitatively examining the consequences of this for variation in the scaling of metabolic rate. At present, the MLB hypothesis (138, 143), has been tested predominantly using comparative data gleaned from the literature, though intraspecific studies have begun (57), and further experimental tests are underway (D.S. Glazier, personal communication). Such tests are sorely needed, as are manipulative tests of the distribution network geometry and DEB theories (202).

A common and usually insurmountable problem associated with tests of explanations for metabolic scaling has been a reliance on correlational approaches to understand the scaling of physiological traits with body mass. This approach precludes examination of the causal effect of mass on the trait of interest. A potential solution to this problem is the examination of scaling relationships for colonial organisms. The size of colonies can be manipulated experimentally and the consequences of the manipulation for scaling relationships can be examined in light of the explicit predictions of competing theories (e.g., 158, 290, 438).

On the metabolic basis of ecology

A valuable research agenda that arose following the publication of the original distribution network geometry models for metabolic scaling over a decade ago (420, 421) was the development of a metabolic theory of ecology (MTE), which attempts to link the size and temperature dependence of individual metabolic rates to size- and temperature-dependent ecological processes at levels of organization from individuals to the biosphere (42, 373). Such links have been recognized for decades (e.g., 43, 314), and an early example of the possible association between metabolism and ecology arises from the observation that of population energy use ($W \text{ha}^{-1}$) is independent of species mass when estimated as the product of metabolic rate ($W \text{individual}^{-1}$, approximately proportional to $M^{3/4}$) and population density (individuals $\text{ha}^{-1}$, approximately proportional to $M^{-3/4}$) (80, 81).
The most prominent MTE is grounded in the first principles explanation for metabolic scaling of the fractal distribution network geometry model of West, Brown and Enquist (421, 419) (WBE hereafter). The WBE-MTE combines the predictions of their model for metabolic scaling with a prediction of the temperature dependence of metabolic rate based on the kinetics of biochemical reactions (134) to arrive and the fundamental equation of WBE-MTE, which describes the size and temperature dependence of metabolic rate with only one free parameter (42). The fundamental equation of WBE-MTE is deliberately simple, yet has been remarkably successful in explaining some ecological patterns (9, 42, 46, 166, 273, 274, 285). Its failure in other cases (8, 84, 161, 282, 331, 344, 403) may stem, at least in part, from violation of assumptions of the basic model (58), or an imprecise description of the effects of temperature and mass on MR (436). Recent work by the proponents of WBE-MTE acknowledges the existence of such variation (e.g., 86, 87), and an interesting avenue for future work is determination of the extent to which incorporation of this variation into the fundamental equation of WBE-MTE improves its predictive power.

While the WBE-based theory is the most well-known MTE, an alternative theory of individual metabolism (DEB) was proposed over 10 years prior to WBE theory, but based on different principles (406). DEB has also been lauded as having the potential to unite hierarchical levels of biological organization (293), and therefore, also represents a possible mechanistic basis for the putative link between individual metabolism and ecological patterns. That a robust MTE can be built upon two very different foundations highlights the important distinction between the general idea behind a MTE (i.e., the hypothesis that metabolic variation underlies ecological patterns and processes) and the particular mechanistic theory that underlies any given MTE (e.g., DEB or WBE). Indeed, any theory that explains the size and temperature dependence of metabolic rate has the potential to form the basis of a robust MTE. Given this, it is essential that the putative mechanistic basis of any MTE is subjected to rigorous experimental testing, again preferably using tests that can distinguish between alternatives (202).

More generally, descriptions of metabolic scaling relationships can be used as a basis for understanding ecological patterns without invoking mechanistic explanations for the scaling relationships themselves. For example, Humphries and Careau (187) used published scaling relationships for the metabolic cost of transport (355) and the HDL model for heat dissipation capacity (385) to predict the influences of body size, ambient temperature, and mode of activity on the extent to which heat generated as a by-product of activity can substitute for thermogenesis. They predicted that, regardless of activity mode, activity-thermoregulatory heat substitution increases with body size and ambient temperature and propose that this offers a simple, null-model alternative to niche-based interpretations of the macroecology of endotherm metabolism (187). Because activity-thermoregulatory heat substitution is predicted to be more common in cold environments than warm ones, the range of energy expenditures observed for animals in warm environments is predicted to be wider than for animals in cold environments, because the energy expenditure of inactive and active animals is similar in the cold but widely divergent in warmer conditions (187). Such a pattern is observed in data for endotherm FMRs (11, 187, 385), and has been suggested to demonstrate that warm low-latitude environments provide a greater variety of feasible metabolic niches than do cool high-latitude environments, and has therefore, been proposed to be a mechanism contributing to latitudinal diversity gradients (11, 70). In contrast to this niche-based interpretation, the analysis of Humphries and Careau (187), based on simple scaling relationships, demonstrates that the latitudinal gradient of variation in endotherm energy expenditure can arise as a simple consequence of temperature dependent activity-thermoregulatory substitution, and does not require an explanation based on hypotheses concerning the variety of available metabolic niches. Such approaches, based on the scaling of physiological traits, have great potential to interface with the continuing documentation of variation in physiological traits over large geographical and temporal scales (Macrophysiology: 68, 131), and can serve to provide a mechanistic underpinning to the understanding of macroecological patterns and to understand and predict the consequences of global change (65, 67).

A complimentary way of viewing the association between metabolism and ecology is not to view ecological patterns as being dictated by variation in metabolic rate, but to view metabolic rate as being shaped by ecological factors. Such an ecological theory of metabolism, as recently proposed by Glazier and colleagues (138, 143, 144, 204), emphasizes that both intrinsic and extrinsic influences on metabolic rate, and the interaction among them, must be considered in any synthetic theory of the scaling of metabolic rate. Certainly, there are many examples where biotic and abiotic variables influence metabolic rate and the scaling thereof (see 138 for a comprehensive review); recent examples include (i) the hypothesis that lifestyle influences the scaling of metabolic rate in fish because differences in lifestyle are associated with differential investment in energetically expensive tissues as required for swimming during predator-prey interactions (204); (ii) the observation that predation influences the growth rate and the intraspecific scaling of metabolic rate in amphipods (144); (iii) the suggestion that an association between catchment land cover and the scaling of FMR in freshwater crayfish is mediated by among site differences in the availability and quality of food (254); and (iv) the hypothesis that the allometric scaling of metabolic rate in marine bryozoans may mean that populations experiencing high rates of habitat fragmentation have greater energy requirements than populations experiencing lower rates of fragmentation (438). Such efforts to understand how ecology shapes metabolism will benefit from consideration of metabolic theories, because these aim to capture physical constraints on metabolic possibilities, and provide null expectations for scaling relationships.
Conclusion

The history of the quantitative study of allometric scaling spans well over a century of investigation, and has encompassed many traits. Throughout much of this period, scaling exponents relating physiological variables to body mass have been determined using ordinary least-squares linear regression of log-log transformed data. During recent years it has become clear that physiological data often violate the assumptions of such an approach, which is appropriate only when data are independent and conform to a two-parameter power relationship with multiplicative error on the arithmetic scale. However, many traits show significant phylogenetic signal, indicating that a phylogenetic perspective is necessary when analyzing such data. Some traits do not conform to a two-parameter power function with multiplicative error, and alternative models including nonlinear regression of untransformed data, three-parameter power functions, and polynomial relationships should be considered for such data. Such approaches have been valuable in describing the scaling of metabolic rate, which remains one of the most widely investigated scaling relationships. The scaling exponent of metabolic rate varies between approximately 0.5 and 1, and varies with activity level as well as between endotherms and ectotherms. A number of explanations for the allometric scaling of metabolic rate have been proposed, but none has gained widespread acceptance. Choosing between these explanations is difficult, because they generally predict similar values of the scaling exponent, and manipulative experiments that test among the unique predictions of the explanations are sorely required.

Acknowledgements

Many colleagues and collaborators have contributed to our work on the scaling of physiological variables and we are grateful to them for the many discussions that have contributed to this review. We also thank Doug Glazier, four anonymous reviewers, and Ted Garland for comments on an earlier version of the manuscript, and the Australian Research Council for funding.

References


344. Rombouts I, Beaurang G, Ibáñez F, Chibana S, Legendre L. Marine cope- 
dod diversity patterns and the metabolic theory of ecology. Oecologia 
345. Rowland JM, Emlen DJ. Two thresholds, three male forms result in 
346. Rubner M. Über den Einfluss der Körpergröse auf Stoff- und 
347. Runciman S, Seymour RS, Baudinette RV, Pearson JT. An allometric 
study of lung morphology during development in the Australian pelican, 
Pelecanus conspicillatus, from embryo to adult. J Anat 207: 365-380, 
2005.
348. Ruxton GD, Houston DC. Could Tyrannosaurus rex have been a scav- 
349. Savage VM, Deeds EJ, Fontana W. Sizing up allometric scaling theory. 
350. Savage VM, Gillingham FP, Woodruff WH, West GB, Allen AP, Enquist 
BJ, Brown JH. The predominance of quarter-power scaling in biology. 
351. Schimpf NG, Matthews PGD, White CR. Cockroaches that exchange 
metabolic energy are continuously survive food and water restriction. 
352. Schimpf NG, Matthews PGD, White CR. Standard metabolic rate is 
associated with gestation duration, but not clutch size, in speckled 
353. Schlader ZJ, Stannard SR, Mündel T. Human thermoregulatory behav-
ior during rest and exercise—A prospective review. Physiol Behav 99: 
269-275, 2010.
354. Schmidt-Jensen K. Field energetics of reptiles and dinosaurs. 
356. Schmidt-Nielsen K. Scaling: Why is Animal Size so Important? 
357. Schmidt-Nielsen K. Field energetics of reptiles and dinosaurs. 
358. Schmidt-Nielsen K. Scaling: Why is Animal Size so Important? 
359. Schmid J, Andersen NA, Speakman JR, Nicol SC. Field energetics of 
free-living, lactating and non-lactating echidnas (Tachyglossus acaule-
360. Schlader ZJ, Stannard SR, Mündel T. Human thermoregulatory behav-
ior during rest and exercise—A prospective review. Physiol Behav 99: 
269-275, 2010.
361. Seymour RS, Baudinette RV, Pearson JT. An allometric 
study of lung morphology during development in the Australian pelican, 
Pelecanus conspicillatus, from embryo to adult. J Anat 207: 365-380, 
2005.
362. Smith RJ. Use and misuse of the reduced major axis for line-fitting. 
363. Smith RJ. Allometric scaling in comparative biology: Problems of con-
365. Smith RJ. Use and misuse of the reduced major axis for line-fitting. 
366. Seymour RS, Lillywhite HB. Hearts, neck posture and metabolic inten-
367. Seymour RS. Raising the sauropod neck: It costs more to get less. 
368. Snell O. Die Abhängigkeit des Hirngewichtes von dem Körpergewicht 
und den geistigen Fähigkeiten. Arch Psychiat Nervenk 23: 436-446, 
1890.
369. Snell O. Die Abhängigkeit des Hirngewichtes von dem Körpergewicht 
und den geistigen Fähigkeiten. Arch Psychiat Nervenk 23: 436-446, 
1890.
370. Snelling EP, Seymour RS, Matthews PGD, Runciman S, White CR. 
Scaling of resting and maximum hopping metabolic rate throughout 
the life cycle of the locust Locusta migratoria. J Exp Biol 214: 3218- 
3224, 2011.
373. Speckmann EJ. Size and shape: The developmental regulation of static allometry in insects. 
375. Speckmann EJ. Size and shape: The developmental regulation of static allometry in insects. 
376. Speckmann EJ. Size and shape: The developmental regulation of static allometry in insects. 


