Chapter 2
Estimation of Body Size in Fossil Mammals

Samantha S. B. Hopkins

Abstract Body mass is a fundamental ecological parameter of mammals with implications for a variety of other ecological characteristics. While it cannot be directly measured in fossil taxa, it can be inferred using allometric relationships between skeletal dimensions and mass derived from extant species. Many such relationships have been described, primarily for dental and limb dimensions. Methods of statistical analysis vary widely, however, in ways with substantial implications for the inferred masses of fossil species. The subset of extant species from which the relationship is derived must be representative of the evolutionary and ecological scope of the fossil taxa for which mass is to be estimated. Increasing computing power and an explosion of phylogenetic comparative methods offer the opportunity to gain an understanding of the processes driving these important empirical relationships.

Keywords Body mass • Proxy • Mammalia • Regression • Allometry • Paleoecology

Introduction

Body size is an essential characteristic of animals, with substantial implications for a number of other ecological and evolutionary parameters (Eisenberg 1981; McMahon and Bonner 1983; Peters 1983; Calder 1984; Schmidt-Nielsen 1984; LaBarbera 1989; Blackburn and Gaston 1994). Size is measured in a variety of ways; among animals it is commonly described in terms of either body length or mass, as these parameters offer the greatest predictive value for ecology. Both of these characteristics translate into the constraints on ecology and evolutionary history, from life history to trophic strategy, evolutionary rates to preferred habitats. As a result, reconstructing body size from extinct organisms is an essential step in understanding their ecology.

Among mammals, body size is commonly described in terms of mass; body length is primarily used for mammals too large to be easily weighed, such as whales (Lockyer 1976), although it has been used for other taxa as well (Iskjaer et al. 1989). Body mass does vary significantly through the lifetime of an organism in response to health, lactation, pregnancy, age, food availability, and other factors (Lindstedt and Boyce 1985; Schulte-Hostedde et al. 2005; Toigo et al. 2006), but paleontologists are generally concerned primarily with estimating average adult size for a species. While mass is easy to measure in living mammal species, requiring only a spring scale for the vast majority of mammals, it cannot be directly measured in extinct organisms. There are, however, a variety of skeletal and dental proxies available for the estimation of body mass. All of these proxies depend on isometric or allometric relationships between various skeletal dimensions and body mass.

The most commonly-applied proxies for mass make use of either dental dimensions (e.g., Legendre 1986; Hopkins 2008; Freudenthal and Martín-Suárez 2013) or the diameters of limb bones (e.g., Scott 1983; Gingerich 1990; Rafferty et al. 1995). Ruff (1988, 2003; Ruff et al. 1991) has also had success estimating mass with the articular dimensions of limb bones. Dental proxies, most commonly used for mammals, take advantage of the relationship between food-processing capacity and metabolism (which is driven primarily by body mass; McNab 1988; Whittaker 1999) and benefit from using the most identifiable and preservable elements in the mammalian skeleton, the teeth. Limb proxies, on the other hand, have the advantage of relying on the more direct relationship between body mass and the load borne by the limbs as they support the body on land. In the case of aquatic mammals, for which neither of these proxies may be effective, other proxies such as cranial dimensions

---

S. S. B. Hopkins (✉)
Department of Geological Sciences, Clark Honors College,
1293 University of Oregon, Eugene, OR 97403-1293, USA
e-mail: shopkins@uoregon.edu

D. A. Croft et al. (eds.), *Methods in Paleoecology: Reconstructing Cenozoic Terrestrial Environments and Ecological Communities*, Vertebrate Paleobiology and Paleoanthropology, https://doi.org/10.1007/978-3-319-94265-0_2

and estimated body volume are necessary. Multiple proxies are often necessary for a robust estimate of mass, and a variety of approaches have been applied to determining which of the various mass estimates are most reliable.

While the allometric relationships between mass and skeletal dimensions have been studied for more than 40 years, there remain substantial problems still to be solved in body mass estimation. In particular, there is a need to confront the difficulties created by using regressions on log-transformed data for estimation. Another factor that complicates mass estimation is the lack of a consistent physical constraint on morphology that can be generalized across all mammals; it is nearly impossible to use the same proxy when studying a diverse assemblage of mammals, as the relationships between mass and skeletal proxies change across mammalian phylogeny. Reconstruction of the body mass of extinct mammals requires careful consideration of phylogenetic constraints, locomotion, functional morphology, and the statistics of regression analysis, prediction intervals, and error propagation. Nonetheless, despite the challenges in the approaches available to us today, there is reason for hope that new computational approaches will solve or at least control for existing problems, and even the strategies currently available can generate ecologically meaningful information about fossil ecosystems and organisms.

Terms

Accuracy: The degree to which an estimated value approximates the actual value. In the estimation of mass, this is probably the primary basis with which to judge the value of a method for paleoecological reconstruction. This is in contrast with Precision; see below.

Allometric scaling: A description of the relationship of overall size to morphology. As size changes, some aspects of morphology change faster or slower than others, leading to changes in shape with changing size.

Body size: A measure of the total size of an organism. Common measures of size in vertebrates include mass, total length, and snout-vent length.

Body mass: The total mass of a live or recently dead organism, conventionally measured as weight with a spring scale.

Geometric scaling: A scaling relationship in which two characters are related mathematically as predicted by their geometry. That is, a volumetric measurement (such as mass) should be proportional to the cube of length (with or without coefficients arithmetically modifying the slope or intercept of the relationship).

Heteroscedastic: Non-normally distributed and skewed. Commonly used with reference to the shape of a distribution of variables.

Isometric scaling: A scaling relationship in which two characters of interest retain the same proportions with size change, resulting in retention of consistent shape over a range of sizes.

Jackknifing: The statistical practice of establishing the predictive value of a regression by removing the data points one by one, analyzing the relationship between the x and y variables, and then using that regression to predict the y-value of the excluded datum. Iterated over the entire dataset, this approach generates an estimate of the capacity of the regression to predict values for new data.

Metabolic scaling: A scaling relationship predicted by the mathematical relationship between size and metabolic rate. The metabolic theory of ecology (Brown et al. 2004) predicts that metabolism scales with mass$^{3/4}$. Hence, other features, such as tooth area, used in fueling metabolism are expected to scale with mass$^{3/4}$. This is in opposition to volumetric scaling, which would predict area scaling with mass$^{2/3}$.

Multiple regression: Linear regression that simultaneously considers multiple independent variables to account for the correlations among those variables in estimating the strength of their linear relationships with the dependent variable. The regression equation takes the form $y = m_1x_1 + m_2x_2 + m_3x_3 + \cdots + b$, rather than $y = mx + b$ as in an ordinary linear regression.

Overfitting: A statistical error in which a regression attributes more of the variation in the dependent variable to the independent variable(s) than is actually causally driven as a result of violations of the distribution assumptions of the statistical method.

Power law regression: A method of regression analysis that, rather than fitting a linear relationship between the independent and dependent variables, looks for the exponential relationship that best explains the distribution of values; the regression equation takes the form $y = ax^m$ rather than $y = mx + b$ as in an ordinary linear regression.

Precision: The uncertainty in an estimated value; the exactitude to which an estimated value is articulated. In its simplest terms, this could be the number of decimal places to which the value is estimated. Note that precision is not the same as accuracy, defined above. A mass estimate value can be highly precise, in terms of being highly replicable and not particularly variable in magnitude, but also quite inaccurate, in being quite distant from the actual mass of the animal.

Rectangular area: The estimated surface area of a roughly quadrangular feature (for the purposes of this chapter, tooth surface area) estimated by multiplying length by width.

Standard error of the estimate: A measurement of the uncertainty in a dependent variable value estimated using a regression. It is commonly calculated as $\sigma_{est} = \sqrt{\frac{\sum (Y-Y')^2}{N}}$

Training data: The data used to generate a predictive statistical relationship. In the case of body mass estimation, these are commonly derived from skeletal measurements of museum specimens of extant species associated with individual body masses. The allometric equations produced from the training data can then be used to estimate body mass where only skeletal dimensions are known.

Volumetric estimation: An approach to body mass estimation wherein the 3-dimensional outline of an organism is reconstructed from its skeleton, and that volume is used to infer mass given either an assumed density or an allometric relationship between volume and mass.

Theoretical/Historical Background

Current approaches to the estimation of body size in extinct mammals trace their roots to a paper by Gould (1975) in which he pointed out that the area of mammalian postcanine teeth scales against mass with positive allometry, likely as a result of metabolism and/or the changes in dietary strategies with increased habitat grain. This work provided a basis for straightforward reconstruction of body mass in any fossil mammal known from teeth. While Gould's descriptions of these allometric relationships for a few key mammalian clades were foundational, subsequent work discussed the drivers of variation in the relationship between size and dental dimensions (Gingerich et al. 1982; Fortelius 1990; Smith 2002). Certainly, the assumption underlying the use of allometric relationships derived from extant mammals to infer mass in extinct species is that the processes shaping those relationships in living species were the same as those in place when the fossil of interest evolved. Whether or not this assumption is true depends on whether our explanations for this allometric relationship are correct (Fortelius 1990).

Many authors have suggested that the relationship between tooth size and body mass is constrained by the energetic demands of size and their relationship to food processing needs (e.g., Gingerich et al. 1982; Fortelius 1985). If this is true, then the fossil taxa for which we reconstruct mass must be similar in physiology (so that energetic demands scale similarly with size), the types of food consumed, and digestive strategy to the range of extant mammals used to produce the allometric curve from which fossil mass is inferred. Tooth size-body mass allometry was proposed by Gould (1975) to represent ¾ power metabolic scaling and hence presumably be similar across all crown group mammals. For a discussion of the complexities of metabolic scaling in mammal teeth, see Fortelius (1990). It has also been proposed that the relationship between tooth area (which is a function of length squared) and body mass (which is directly related to volume, a function of length cubed) should show ⅔ power geometric scaling (Gingerich et al. 1982). There is enough noise in the relationship between mass and tooth area (as a result of measurement error, individual variation in tooth size and mass, and interspecific variation in the relationship between the two), and enough different ways of measuring tooth dimensions that the nature of the scaling relationship remains controversial (a problem that plagues the geometric/metabolic scaling problem in general; White et al. 2009). As a consequence, most approaches to mass reconstruction are empirically-based rather than having a strong theoretical foundation.

While early efforts used the area of the entire postcanine tooth row to infer allometric relationships with mass (Gould 1975), Gingerich et al. (1982) and several subsequent authors (Legendre 1986; Van Valkenburgh 1990; Gordon 2003) demonstrated a strong relationship between body mass and the areas of individual teeth, in particular the first molar. The first molar is particularly useful because it tends to have relatively low levels of intraspecific variation and lower levels of sexual dimorphism, likely as a result of its eruption early in ontogeny, prior to puberty (Gingerich 1974). Inferring mass from individual teeth adds to the metabolic scaling assumption by assuming a consistent relationship between the size of that tooth and that of the rest of the tooth row. This assumption can often be assumed to hold true where the dental formula and gross dental morphology remain similar but becomes more questionable with changes in the functional role of the individual tooth within the tooth row. Changes in diet that emphasize the function of particular teeth cause increases or decreases in the relative sizes of teeth, changing these allometric relationships. A widely-used set of regression equations of lower first molar (m1) rectangular area (that is, length × width, rather than the actual area measured within the tooth outline) against body mass for a variety of mammalian clades was developed by Legendre (1986), which led to the adoption of m1 area as the most common proxy for mammalian body size. This proxy has the additional advantage that the lower first molar of the vast majority of mammalian species is among the most taxonomically diagnostic of skeletal remains.

The allometric scaling of limb bones with mass has seen a great deal of work (e.g., Gingerich 1990; Biknevicius 1999; Christiansen 2002), but one key realization is that the load-bearing strength of a columnar limb is proportional to its cross-sectional area (Ruff 1990), so one can use the cross-section of limb bones to infer mass in extinct mammals. This method was shown to predict mass much more reliably than dental dimensions in an array of studies that arose out of a workshop on body mass held at the University of Florida (Damuth and MacFadden 1990), as well as some subsequent work (e.g., Egi 2001). However, this approach has a different set of assumptions that can be violated by extinct mammals. For example, using humeral cross-sectional area as a predictor of body mass assumes similar posture to the modern dataset from which the fossil species' mass was inferred. This has been problematic in some past studies (Sánchez-Villagra et al. 2003; Millien and Bovy 2010; Basu et al. 2016), which have concluded that weight-bearing posture differed in the extinct species under study. Other studies have shown that this scaling relationship may change over the range of extant mammals (Biewener 1990; Christiansen 2002) as the physical constraints change in importance allometrically, so inferences should be made with reference to species within the size range of extant comparators. Ruff (2003) showed that the distribution of bone within the shaft (measured as the section modulus) is more effective at predicting mass than the breadth of the bone shaft, so where precision is important and fossil material is well-preserved, this more labor-intensive approach may yield greater accuracy.

Breadths and areas of limb bone articular surfaces provide another functionally constrained parameter; while they may be affected by locomotor behavior (Ruff 1990, 2002), there is some evidence (Ruff 1988, 1990, 2003) to suggest that they may be more reliable than many other proxies for mammalian body mass. The articular surface area is assumed to be constrained by mass in that excessively large forces exerted on articular surfaces in the joints increase the chances of damage to cartilage and soft tissue. These methods require preservation of identifiable limb elements with complete articular surfaces, but have proven to predict mass well in primates (Ruff 1990, 2003). They have not yet been widely applied among other mammalian groups.

Recently, there has been a dramatic increase in the application of phylogenetic comparative methods to paleoecology. Given the limitations mentioned above to body mass estimation, it is worth asking whether evolutionary history has a substantial influence on the outcome of body mass reconstructions, given that these methods are often applied to fossil taxa with no close living relatives. Campione and Evans (2012) determined that limb bone circumferences do show some component of phylogenetic signal in the fit between limb dimensions and mass but concluded that it was not significant for reconstructions of ecology. No study has made a comprehensive effort to examine the influence of phylogeny in dental reconstructions of body mass, but studies at the ordinal level within mammals (Legendre 1986), between families of primates (Kay and Ungar 1997; Scott 2011), and within families and superfamilies of rodents (Lindsay 1988; Morgan et al. 1995; Hopkins 2008) have shown that regression equations can differ substantially between clades, suggesting the importance of phylogeny in constraining these relationships.

Approaches

Given the variety of proxies available, any study of fossil mammal body mass has a number of methodological options. Choice of an appropriate proxy for body mass reconstruction should be dictated by the biology of the group under study. Where possible, multiple proxies should be used and compared to determine which estimate is most likely to be correct and the range of possible values for body mass. While one can obtain a very precise (but not necessarily very accurate) estimate of mass from previously-published allometric regressions, even a very good mass proxy generally has notable (though often poorly reported) uncertainty, and an honest reconstruction of mass for a fossil mammal includes an assessment of this uncertainty.

The methodology is similar for almost all proxies commonly used. Data are collected on the proxy measurements and body mass from a training dataset of extant species, commonly from mammalogy collections, which often have live mass data for many of their museum specimens. There is disagreement about whether or not it is necessary to collect data from a sample of individuals for each species or not; the argument has been made (Legendre 1986; Hopkins 2008) that the goal is a species average, so individual variation isn't essential to the outcome. However, if the single individual measured is unusual, it can be misleading, especially if the training dataset is small. Even a small sample of 3–5 individuals can increase the robustness of estimates. Values for the sample are often averaged to generate a data point for the species. Many studies use individual masses from the specimens measured (e.g., Reynolds 2002; Hopkins 2008), but others simply use published average masses for the species (e.g., Gingerich et al. 1982; Janis 1990; Mendoza et al. 2006) or a mix of published and individual masses (e.g., Ruff 1990; Freudenthal and Martín-Suárez 2013). Published values are adequate when necessary, but the correspondence is likely to be less accurate; the individuals sampled may or may not be representative of the species average for dental measurements, and a mismatch could lead to an inaccurate placement in the allometric equation. The species data, both mass and proxy, are generally log transformed (either natural log or base 10), and then a linear regression is applied to establish the allometric relationship between the proxy and mass in log-linear space (Fig. 2.1). For some fossil taxon of interest, the proxy is then measured, and the regression equation can be used to infer mass.

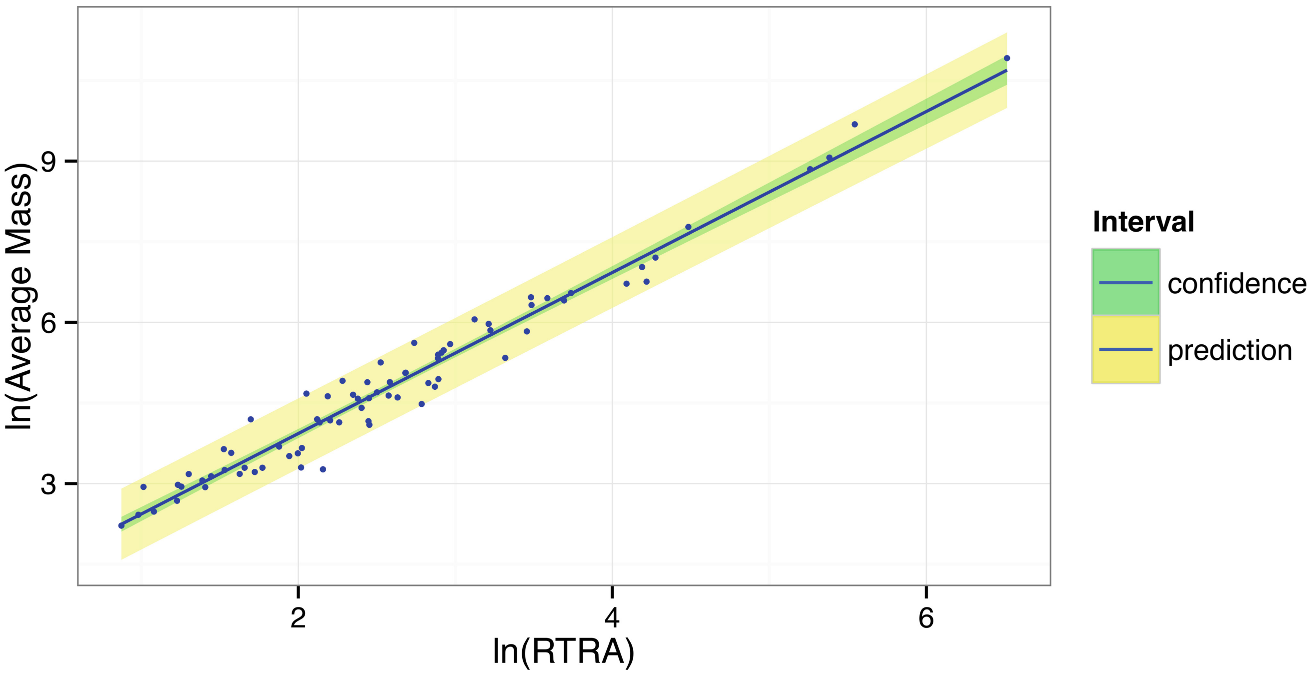

Fig. 2.1 Example regression in ln-ln space of a mass proxy (in this case RTRA, or lower cheek tooth row area) against body mass for a training dataset of 76 extant rodents. Data from Hopkins (2008). Points represent species average values. The green shading shows the 95% confidence interval for the regression line, while the yellow shading represents the 95% prediction interval

Examples of widely-used studies generating mass estimation equations are referenced in Table 2.1 with information about the taxonomic group and morphological proxy or proxies they use as well as the methods applied. The regression equations can be found in the papers, along with information about the training data used to generate them.

Dental Proxies

The most common proxy used to estimate the body mass of fossil mammals is the size of the first lower molar, m1 (Alroy 1998; Gingerich et al. 1982; Legendre 1986; Van Valkenburgh et al. 2004). This proxy is often measured in terms of area, approximated by multiplying the maximum mesio-distal length by maximum bucco-lingual width, assuming a roughly rectangular occlusal shape (Fig. 2.2). Other authors have found length alone to be adequate and, in fact, to be less sensitive to dietary differences when estimating mass of large mammals (e.g., Damuth 1990). Regressions using this proxy are available for a variety of mammalian groups (Legendre 1986) including ungulates (Janis 1990; Mendoza et al. 2006), marsupials (Gordon 2003), carnivorans (Van Valkenburgh 1990), and especially for primates (Gingerich et al. 1982; Conroy 1987; Dagosto and Terranova 1992; Vinyard and Hanna 2005; Copes and Schwartz 2010; Scott 2011). The m1 has the least variation in its fit to body mass of all the cheek teeth (Gingerich et al. 1982; Janis 1990; Gordon 2003) and therefore should be the most precise single tooth dental proxy for reconstructing body mass, though regressions are available for most of the other teeth in the dentition as well (see Table 2.1 for examples).

Because dental formula varies among mammals, another proxy uses the area of the entire postcanine tooth row for body mass reconstruction, taking advantage of the partitioning of function among the mammalian teeth (see also Evans and Pineda-Munoz 2018). As for a single tooth proxy, this value is most commonly estimated by multiplying the maximum anteroposterior length of the postcanine tooth row by the maximum width of the cheek teeth (Gould 1975; Hopkins 2008; Fig. 2.2). Note that this proxy would not be expected to work as well as single-tooth proxies in taxa with substantial curvature to the postcanine dental arcade, as occurs in some primates. Tooth row length can be used similarly (Hopkins 2008) but only in a clade for which the proportions and numbers of teeth are relatively constant. Both of these proxies are useful for taxa such as muroid or mylagaulid rodents with unusual first molars but require a fossil record that preserves undistorted tooth rows. Tooth row area and length can also be approximated from the sum of isolated cheek teeth (Freudenthal and Martín-Suárez 2013), although given mesio-distal overlap of some cheek teeth, this method will necessarily introduce some noise into the estimates of mass.

Table 2.1 Example body mass proxy studies. This is not intended to be a comprehensive list; rather, it is intended to show some of the diversity of available proxies, particularly those that are widely used or methodologically novel. Abbreviations in methods: MLR = multiple linear regression; OLS = ordinary least squares; PIC = phylogenetically independent contrasts; RMA = reduced major axis

Reference	Type of proxy	Proxy measurement(s)	Regression method	Taxonomic scope
Gingerich et al. (1982)	Dental	Area of each tooth of the tooth row; Cheek tooth row area	Ln transform, OLS	Primates
Anderson et al. (1985)	Postcranial	Total stylopodial circumference	Log$_{10}$ transform, OLS	Ungulates
Legendre (1986)	Dental	m1 area	Ln transform, OLS	Mammals, Chiroptera, Eulipotyphla, Carnivora, Primates, Rodentia, Ungulates
Damuth (1990)	Dental, Whole body	Head-body length, Length, width, and area of individual teeth, molar row, and cheek tooth row	Log$_{10}$ transform, OLS, MLR	Ungulates
Gingerich (1990)	Postcranial, Multivariate	Length and diameter of femur, humerus, tibia, metacarpal, and metatarsal; Ulna length	Log$_{10}$ transform, OLS, MLR	Mammalia
Janis (1990)	Dental, Cranial	Lengths, widths, and areas of each individual postcanine tooth; Premolar row length; Molar row length; Muzzle width; Masseteric fossa length; Occipital height; Posterior skull length; Basicranial length; Anterior jaw length; Posterior jaw length; Depth of mandibular angle; Width of mandibular angle; Jaw length	Ln transform, OLS	Artiodactyla, Perissodactyla, Procaviidae (Hyracoidea), Macropodidae (Diprotodontia)
Martin (1990)	Dental	m1 area	Ln transform, OLS	Cricetinae
Ruff (1990)	Postcranial	Length and breadth of femur and tibia; Cross-sectional shape and area of femur and tibia; Proximal and distal articular surface areas of tibia and femur	Log$_{10}$ transform, OLS	Anthropoidea (Primates)
Van Valkenburgh (1990)	Dental, Cranial, Whole body	Head-body length; Skull length; Occiput-to-orbit length; m1 area	Log$_{10}$ transform, OLS	Carnivora, Dasyuridae (Dasyuromorphia)
Reynolds (2002)	Cranial, Postcranial	Condylobasal length; Skull length; Femur length	Log$_{10}$ transform, OLS and RMA	Beavers and giant beaver analogs
Mendoza et al. (2006)	Cranial, Dental, Multivariate	Lower tooth row length; Premolar tooth row length; Lengths and widths of individual postcanine teeth; Anterior jaw length; Posterior jaw length; Depth of mandibular angle; Width of mandibular angle; Length of coronoid process; Length of masseteric ridge; Occipital height; Posterior length of skull;	Log transform, MLR, phylogenetic weighting	Artiodactyla, Perissodactyla, Procaviidae (Hyracoidea)

(continued)

Table 2.1 (continued)

Reference	Type of proxy	Proxy measurement(s)	Regression method	Taxonomic scope
		Depth of face under orbit; Length of paraoccipital process; Muzzle width; Palatal width; Basicranial width		
Hopkins (2008)	Dental	Tooth row area; Tooth row length	Ln transform, OLS	Rodentia, and several families therein
Millien and Bovy (2010)	Dental, Postcranial, Multivariate	Skull length; Upper cheek tooth row; Upper and lower incisor length and width; m1 length; m1 width; Humerus length; Femur length; Humerus diameter; Femur diameter	Log_{10} transform, OLS and MLR	Caviomorpha (Rodentia)
Campione and Evans (2012)	Postcranial	Total stylopodial circumference	Log_{10} transform, OLS, PIC	Mammalia, Ungulates, Carnivora, Marsupialia, Euarchonta, Glires
Freudenthal and Martín-Suárez (2013)	Dental	Tooth row area	Ln transform, OLS	Rodentia, and several families therein

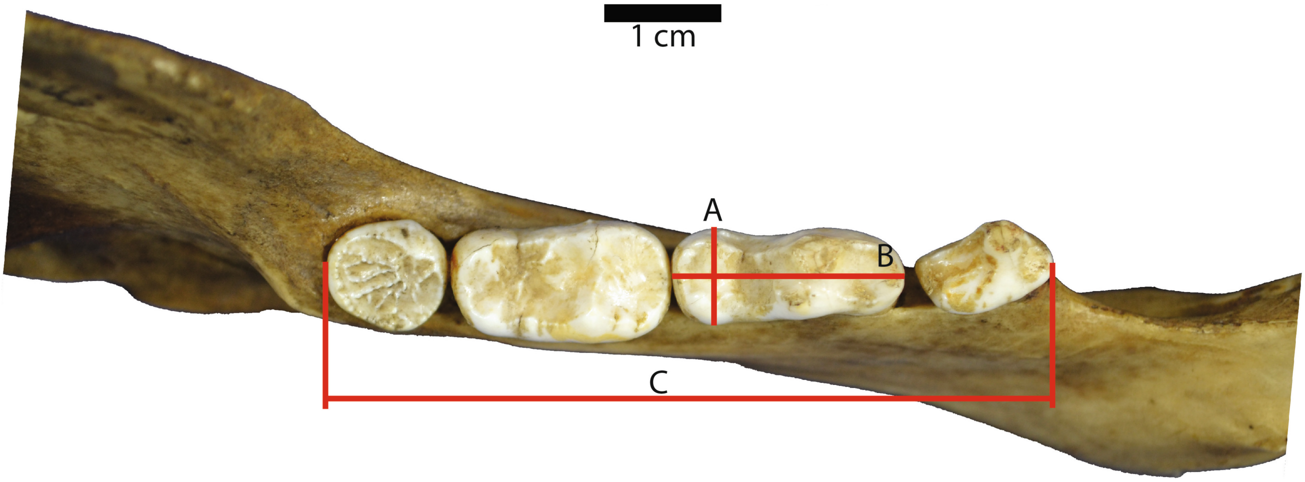

Fig. 2.2 Examples of three dental dimensions commonly used for mass regressions, illustrated on the lower tooth row of a black bear (*Ursus americanus*), University of California Museum of Vertebrate Zoology mammal specimen #8816. **A** = m1 width; **B** = m1 length; **C** = lower tooth row length

Limb Bone Proxies

Limb-based mass estimation of body mass uses a greater diversity of proxies than dental estimation. The most commonly-employed proxies use either lengths of long bones (Gingerich 1990; Reynolds 2002) or the midshaft cross-sectional dimensions of the humerus (Gingerich 1990; Millien and Bovy 2010), femur (Ruff 1990; Reynolds 2002), or both (Anderson 1985). Bone lengths are generally measured with calipers or measuring tapes (Gingerich 1990; Reynolds 2002); this proxy is useful because it is also commonly published in taxonomic descriptions and can be measured from figures in publications with reasonable accuracy. In contrast, cross-sectional dimensions (including area, perimeter, and diameter) usually require direct measurement of museum specimens; these values are seldom published except in studies of body mass estimation, and only rarely are published images sufficiently accurate to allow measurement of even diameter. In some cases, these proxies are measured digitally from bone scans (Ruff 1990), but they are also commonly measured directly with a measuring tape (Campione and Evans 2012). Limb bone cross-section as a mass proxy depends on the physical relationship between the load-bearing capacity of a cylinder and its cross-sectional area (Campione and Evans 2012), and as long as the work adequately addresses postural differences among species in the training data (Sánchez-Villagra et al. 2003; Basu et al. 2016), these proxies can be among the most universally applicable (Campione and Evans 2012). Long bone lengths require a tight morphological, functional, and phylogenetic match to provide reliable mass estimates; good analogs for locomotor habits of fossil taxa must be used and, as pointed out by Biewener (1990) and Christiansen (2002), large mammals (above 50 kg) have relatively shorter limb bones for their mass than smaller species due to a shift in the primary constraint on limb shape from compressional stresses to bending strain. Because limb-based estimates have been used most often for larger mammals (in part because identifying postcrania of large mammals is often easier), many of which are above this threshold, this change is less problematic than it would be if the proxy were applied more widely. This limitation should be an important consideration in any study reconstructing a wide range of masses.

While limb shaft dimensions are the most widely used proxy for mass estimation, the ends of the limb bones can also be used. Articular surface dimensions, often as linear or approximate area measures, can offer an alternative approach (Ruff 1988, 1990, 2003). However, their reliability as a mass proxy is somewhat contingent on posture (Ruff 1990, 2002), and at least one study (Ruff 2003) found that cortical bone distribution (obtained from CT scans of limb bones) is a better predictor of body mass that external dimensions. In a case where higher precision is required, these more labor-intensive approaches can increase the robustness of the physical relationship on which the inference depends.

Volumetric Estimation

In taxa with no good living analogs, several authors (Bates et al. 2015; Brassey et al. 2015; Basu et al. 2016) have argued that the best method of body mass estimation is to make a complete skeletal reconstruction, overlay soft tissue on that reconstruction, and estimate body mass based on volume and estimated average density. Such an approach enables reconstruction of mass in some taxa for which other proxies fail but requires exceptionally complete fossil preservation and substantial modeling effort. Hence, it is not efficient or even possible for many applications. It has, however, allowed reconstruction of mass in animals such as dinosaurs (Bates et al. 2015; Brassey et al. 2015) and the extinct giraffe relative *Sivatherium* (Basu et al. 2016) for which simpler mass proxies are suspect.

Considerations

Available fossil record: A decision about mass estimation methods begins with the skeletal elements preserved and well-identified for the taxon of interest. If postcrania cannot be confidently identified to species, dental estimates must be preferred in spite of their decreased accuracy. If the study requires reconstructing mass of a large number of distantly-related species, it may be necessary either to use multiple proxies or to choose a proxy that does not necessarily reconstruct mass with the greatest accuracy or precision but that does allow all the species of interest to be included or minimizes the violations of regression assumptions for taxa in the analysis.

Physical constraints: Regression equations only apply where the training data are representative of the species for which mass is being estimated. For example, it is challenging to reconstruct mass for taxa well outside the size range of extant relatives such as the extinct giant rodents *Phoberomys* and *Josephoartigasia* (Hopkins 2008; Rinderknecht and Blanco 2008; Millien and Bovy 2010), as the physical constraints that shape the relationship may change between extant and extinct species (Millien and Bovy 2010; Basu et al. 2016). In these cases, it can be more accurate to use more distantly related but more physically similar species as training data.

Dentition: In using dental proxies for body mass, it is necessary to consider how the dentition compares with that of the taxa from which the regression equation was derived. Taxa with dramatically different diets tend to have somewhat different constraints on the shape and scaling of the cheek teeth (Gingerich et al. 1982; Legendre 1986; Janis 1990; Scott 2011; Evans and Pineda-Munoz 2018). Regressions based on training data including taxa with similar dietary ecology are most likely to reconstruct mass accurately. There are even a few ecologies explored by mammals (for example, colonial insect specialists) that release the constraint on dentition size because teeth are not used in food processing. In such a case, dental proxies are not useful or cannot even be used for mass reconstruction.

Locomotion: In using proxies based on limb bone dimensions, we make the assumption that the limbs are loaded in the same way as those in the training data. If the distribution of weight or the limb loading in locomotion is dramatically different in study species relative to training data, mass reconstructions will be inaccurate. If the problem is simply distribution of weight (for example, a species that bears more weight on the forelimbs than its extant relatives), combining the limb cross-sectional areas can still yield accurate estimates of mass. However, if there is a difference in locomotor modes, such as a ricochetal (hopping) species whose extant relatives are all terrestrial, reconstructing mass from limbs is much more challenging; the magnitude of forces experienced during locomotion relates to mass quite differently, with the possible result that dental proxies are a better option. Even complex methods that reconstruct 3D body geometry may not be any more accurate in these cases, as reconstructing musculature and body shape requires reference to modern analogs (Carrano and Hutchinson 2002; Hutchinson and Garcia 2002; Basu et al. 2016). Limb posture also dictates which physical constraints are important for limiting limb dimensions, so mammals with unusual limb postures (for example, Mesozoic mammals) should be compared to species with analogous posture where possible.

Evolutionary relatedness: The shape of mammalian teeth, limb posture, locomotor habits, and even body density and muscle mass vary over evolutionary time and show substantial phylogenetic constraints, so it is important to consider evolutionary relatedness in choosing which of the available regression equations to use in reconstructing body mass. Studies of dental regressions, in particular, have demonstrated substantial and significant differences in regression coefficients between different taxa (Legendre 1986; Damuth and MacFadden 1990; Hopkins 2008). There are also some taxa for which these regressions are simply more variable and hence less precise in reconstructing mass. Limb bone dimensions offer a less phylogenetically-constrained proxy that has been argued to be consistent across wide swaths of terrestrial vertebrates (Campione and Evans 2012).

Statistical Methods: For those developing their own mass estimation regressions, a few statistical best practices are worth considering. Choice of regression methods, data transformation, and realistic examination of statistical errors are all very important for generating robust mass estimates for fossil mammals. Existing regressions differ widely in the degree to which these recommendations have been followed; it is worthwhile to consider what effects these methodological differences could have on the resulting mass estimates.

Regression methods: The most commonly used method for linear regression is Ordinary Least Squares (OLS) Regression. While some authors have argued (Ricker 1973, 1984; Reynolds 2002) that a different method, such as Major Axis or Reduced Major Axis Regression, is more appropriate given that they allow for equal levels of uncertainty in the x and y axes, this argument applies primarily to determining the strength and direction of allometric relationships and not to mass prediction. OLS regression assumes that the independent variable (in this case the proxy measurement) is known absolutely, while the dependent variable contains all the error in the system (Zar 2010); this is true for body mass estimation, because the measurement of the proxy is generally accurate to within a very small error, while the mass is generally much less certain, given the issues discussed previously.

Multiple regression methods have sometimes been applied in cases where multiple skeletal elements are available for a given taxon (Gingerich 1990; Mendoza et al. 2006; Millien and Bovy 2010). While it seems intuitive that adding information from multiple proxies would improve the accuracy of mass estimates, this is not always true (Kaufman and Smith 2002); unless the number of species sampled is large, there are problems in overfitting when the ratio of sampled species to predictors (that is proxy measurements) is less than 10 (Smith 2002; Kaufman and Smith 2002). Multiple regression is also not appropriate to selecting the best proxy from among a variety of candidate mass proxies, because the multicollinearity of the proxies makes such a process unreliable (Kaufman and Smith 2002). The practice of generating multiple bivariate regressions and averaging the resulting mass estimates (statistically an unjustifiable approach) does not consistently escape the problems of multiple regression and involves an additional assumption about the independence of measurements that rarely holds true (Smith 2002). In some cases, averaging proxies improves the answer where the errors of the different estimates average out (Scott 1990), but because this relies on the included proxies erring roughly equally in either direction of the true mass, it cannot be relied upon consistently. Instead, it is generally better to choose the most reliable univariate proxy based on the biology of the taxon of interest or to choose a small number of independent proxies for multivariate analysis (Kaufman and Smith 2002).

Data transformation: Size data in living systems often have a heteroscedastic distribution of variance because in larger organisms, the magnitude of variation in size is greater. Hence, the residuals from a regression line suggest that the variance is not evenly distributed, as assumed by the vast majority of linear regression methods. In the past, the solution has commonly been to log-transform the values both for the proxy measurements and the masses (see Table 2.1). This does solve the problem of heteroscedastic variance in many cases, although it is necessary to examine regression residuals to determine whether $\log_{10}$ or natural log adequately account for differences in variance (Zar 2010). In some cases, they do not perform equally well, although both methods have been applied to body mass proxy data. Unfortunately, log-transformation of the mass proxy data creates some analytical problems, a topic discussed below (see "Challenges"); there is currently no clear answer about whether log-transformation leads to values that would be more or less reliable than untransformed results (Smith 2002; Kerkhoff and Enquist 2009; Packard 2009, 2013).

Statistical uncertainty: Because these relationships between body mass and skeletal dimensions are empirically determined, they have error. In order to use these regressions meaningfully, it is necessary not only to generate a mean estimate for mass but also to make an honest statement of the uncertainty in that estimate. Sources of error that must be included in that uncertainty include measurement error in the mass proxy, uncertainty in the regression line that relates the proxy to the inferred mass, variance between individuals in a population, and uncertainty in which proxy best predicts mass. The first two points are statistical; the last two are biological. Many studies do a good job of exploring the uncertainty in mass depending on the proxy chosen (e.g., Gingerich et al. 1982; Sánchez-Villagra et al. 2003; Freudenthal and Martín-Suárez 2013). It is also reasonably common to measure a range of specimens and take a mean of that sample before inserting that value into a mass regression (e.g., Janis 1990; Ruff 1990; Delson et al. 2000), although it is somewhat less common for those regressions to include measurements of multiple individuals of the species in the training dataset. Few studies discuss the measurement error in the mass proxy, likely because the error of measurement is assumed to be proportionally much smaller than the error in the prediction equation, although this may be less true in very small taxa. While a regression analysis commonly generates a standard error of the estimate that summarizes the error across all these sources of uncertainty, the error in a mass regression is often not evenly distributed across the range of values; it is commonly greater at the extremes of the distribution. An accurate assessment of uncertainty in the mass estimate for a given taxon depends on where in the distribution its size falls, and hence requires a predictive equation of its own generated through the use of statistical error propagation. It is very rare, though, to properly propagate error through the regression equation (although see Hopkins 2008), even though the error in the x and y dimensions of the original training data are compounded by the error in the x measurement for the fossil species of interest (Zar 2010). It is worth mentioning that biological sources of error may also generate statistically non-normal distributions of variation, which causes many conventional statistical estimates of error, such as standard error, to skew uncertainty inappropriately. There is little research on the role this problem may play in generating inaccurate mass estimates, although it may be important to the persistent skew in prediction errors found by Freudenthal and Martín-Suárez (2013).

One critical dimension of statistical uncertainty that is desperately underreported is the uncertainty in an estimate produced by the equation; this uncertainty varies across the range of sizes encompassed by the training data (with lower uncertainties in the middle of the data range), but reported equations rarely give users the information needed to estimate the uncertainty in a mass prediction generated by a given equation. This is a critical parameter to mass estimates that cannot be replaced by a simple average prediction error value over the whole training dataset.

Empirical uncertainty: Given the challenges unique to the different systems for which these regressions have been developed, it is vital for a regression analysis to generate meaningful error estimates. One of the most valuable strategies is jackknifing (Wu 1986), in which each of the data points in the training dataset is removed one at a time and then predicted using the remaining data. Comparing these predictions with the observed data will make it possible to assess, within the data used for training the regression, how much error is likely in mass estimates and whether there are any systematic biases in the regression result. This helps with both the log-averaging problem and the problem of error estimation presented above, though it is an estimate relevant only to the breadth of the training data, and hence still limited by the degree to which those data are representative of the fossil taxa being reconstructed.

Strengths

Mass is correlated to many other aspects of a mammal's biology, possibly more so than any other biological variable, making it a vitally important aspect of the biology of an extinct mammal. Mass has demonstrated correlations with diet (Price and Hopkins 2015), habitat use (Peters 1983), rates of evolution (Martin and Palumbi 1993; Gillooly et al. 2005; Liow et al. 2008), metabolic rates (McNab 1988; West et al. 1997; Whittaker 1999), life history strategies (Sibly and Brown 2007; Liow et al. 2008, 2009), home range size (McNab 1963; LaBarbera 1989), and a variety of other biological variables that are much more difficult to infer from the fossil record than is mass. Hence, a robust estimate of body size has the capacity to predict other aspects of biology that may not otherwise be available.

Many commonly-used skeletal proxies are very highly correlated with mass, with correlation coefficients generally between 0.95 and 1, yielding surprisingly precise estimates, at least of species average mass values. Mass can be reconstructed from a variety of proxies, depending on the available material, because overall body size is one of the most important predictors of the size of individual skeletal elements. This variety of potential proxies makes it possible to estimate mass for almost any fossil mammal, even those for which preserved remains are extremely rare or fragmentary or those that have evolved unusual limb or tooth morphologies (Biknevicius 1999; Hopkins 2008; Pyenson and Sponberg 2011). Because these methods are non-destructive, they allow reconstruction of a critical ecological dimension even for a species known from a single isolated tooth (although estimates are obviously more robust for a larger sample).

Mass data are also frequently available for extant mammals in museum collections, in ecological publications, and in a variety of databases, making it easy to compare fossil ecological assemblages with modern ones. One of the classic papers in mammalian body mass reconstruction (Legendre 1986) made use of this phenomenon to reconstruct habitat types from body mass distributions among fossil mammal communities, although this inference remains somewhat controversial (Rodriguez 1999). There are certainly some significant ecological patterns in body mass distributions that can be valuable for understanding large-scale ecological processes.

Biases and Shortcomings

While these methods of body mass estimation are powerful, there are also some noteworthy problems with reconstructing body mass in fossil mammals that must be acknowledged. First, we are generally limited to proxies commonly available from the fossil record. While there are essentially limitless possibilities for mass proxies, many of those that may be some of the best predictors of mass may not apply usefully to many taxa of fossil mammals because the elements required are rarely preserved. For example, the breadth of the occipital condyles offers an excellent mass proxy that may be consistent across a wide swath of mammalian taxa with minimal variation even in very different ecologies, sizes, or phylogenetic lineages (Martin 1980; Pyenson and Sponberg 2011). However, because the majority of mammalian taxa are identified using dental features, reconstructing mass with this proxy requires exceptional preservation of an undistorted basicranium with associated dentition, so this proxy is less useful for many mammalian species. Alroy (2012) showed that the size of the glenoid fossa of the scapula is one of the best predictors of mass in mammals; sadly, scapulae are rarely identified to species and almost never preserved intact, so these proxies are of very limited utility in estimating fossil mammal masses in spite of their accuracy and precision. While they may provide a useful check on other proxies in cases of exceptional preservation, they are not applicable to the majority of fossil mammals.

A second shortcoming, as previously mentioned, arises when proxies disagree in their estimate of body mass (Hopkins 2008; Millien and Bovy 2010; Basu et al. 2016). Determining which, if any, of the estimates are correct is a challenge that requires additional study of the paleobiology of the species of interest. The discussion in this chapter of the processes underlying different proxies should provide some basis for such a study, but in the end, this problem makes the simple estimation of this fundamental ecological parameter much more complex in some species.

Phylogeny can bias results in a variety of ways. First, the allometric relationships between various skeletal dimensions and overall size are generally strongest within narrow clades and much less consistent as species diverge evolutionarily, although Freudenthal and Martín-Suárez (2013) argue that these differences are not significant. Within families of mammals and even within orders, regressions are often robust and quite accurate (Legendre 1986; Hopkins 2008; Freudenthal and Martín-Suárez 2013), but morphological constraints (such as dental formulae, locomotor modes, etc.) can change enough between families and orders to render mass estimates from these proxies less reliable (Legendre 1986; Hopkins 2008; Millien and Bovy 2010). In addition, there are some narrow clades for which particular proxies or even all proxies may be less reliable. For example, muroid rodents differ from most other rodents in having lost the fourth premolar in the upper and lower dentition. As a result, the first molar, to which many mass estimates are tied, might be expected to be more variable, as it has become the anteriormost cheek tooth; this position is generally expected to be released from stabilizing selection by the absence of an articulating tooth on the mesial side (Gingerich 1974). Indeed, muroid mass is much less reliably predicted by m1 dimensions than other rodents (Lindsay 1988; Morgan et al. 1995). However, this expected release of constraint on m1 is not consistent with the findings of Freudenthal and Martín-Suárez (2013) wherein variation in muroid teeth was lower in m1 than m2, nor does it explain why the tooth row length and area were also more variable in their relationship to mass in muroids than in any other rodent clade (Hopkins 2008). It appears that the relationship between dental dimensions and mass is simply more variable in this clade, suggesting that this correlation may also be a property that varies phylogenetically. Glirids also present challenges to mass prediction via dental dimensions. Freudenthal and Martín-Suárez (2013) present the hypothesis that differences in digestive efficiency may explain confusing results of fossil and extant glirid body mass estimates. The dentitions and limbs of whales are also poor predictors of mass (Pyenson and Sponberg 2011), likely the result of a similar release of constraint in these aquatic creatures. These issues become particularly important in reconstructing mass for mammals well outside of extant orders and families. A better understanding of the reasons for variation in the relationship between mass and various size proxies will make estimates of mass more robust.

One of the more controversial problems in mass estimation is the log-averaging problem. In short, the use of log transformation to reduce heteroscedasticity in the variance of training data leads to consistent under-prediction of mass by proxy data (Smith 1993, 2002). However, the distribution of error in dental and skeletal mass proxy data is multiplicative (Huxley 1932; McMahon and Bonner 1983; Gingerich 2000; Kerkhoff and Enquist 2009; Xiao et al. 2011; Glazier 2013), and hence requires some kind of transformation to correct for the heteroscedasticity of this error. There are some mathematical solutions to the problem (Smith 1993) that approximate a correction and work effectively in some cases (Myers 2001), but the bias cannot be completely and consistently removed as long as regressions are performed on log-transformed data. The bias is relatively small in magnitude, perhaps 5–15% for most regressions, but gets larger as the range of sizes included in the regression equation increases or as the prediction error increases. Freudenthal and Martín-Suárez (2015) considered the problem when estimating rodent head-body length and found that predictions from log-transformed data were low at the high and low extremes and high in the center of the distribution, from which they concluded that log transformation should be avoided when possible. This problem implies that statistics that are robust to non-normal distributions may be necessary in some cases. Packard (2009, 2013) has argued that regression algorithms are now robust enough that the data transformation may not be necessary and that a power law regression can be run on untransformed data to study such allometric relationships. It remains to be seen whether this solution is applicable to mass regressions; in general, the non-normality of the data in most mass regressions is large enough that standard statistical practice (Zar 2010) would demand data transformation prior to using parametric regression methods. Kerkhoff and Enquist (2009) argue that the geometric averaging inherent to regressions on log data actually yields a more biologically realistic description of variation and that log-transformed data are actually more appropriate for allometric regressions. The question of whether arithmetic regression can be applied to allometric relationships such as this one has not yet been adequately tested in body mass proxy regressions, so most workers continue to apply log transformation to body mass regressions (White et al. 2009; Millien and Bovy 2010; Field et al. 2013; Freudenthal and Martín-Suárez 2013; Brassey et al. 2015).

Examples of Applications

Size Evolution over Deep Time

Our understanding of mass evolution across the mammalian lineage relies heavily on skeletal proxies for mass reconstruction. Smith et al. (2010) used mass estimates for fossil terrestrial mammal species to find the largest species in each mammalian order through the Cenozoic. A study so broad-reaching required application of a variety of dental and limb-based mass proxies (depending on the identified material available for each of the described species) although the effort admittedly was not comprehensive, focusing only on the largest (in size, not diversity) clades in each order. While the use of varied proxy data, including some for which the consistency of the proxies was not assessed, is less than optimal, this study aimed for a broad-brush picture of maximal size by order and for mammals as a whole through time. As a result, the differences in precision and accuracy between the varied proxies were not problematic for the purpose to which the size estimates were applied. Roth (1990, 1992) solved the problem of intercomparability of proxies and of increasing error with size by simply comparing the raw proxy values in her studies of dwarfing in fossil elephants. Gould and MacFadden (2004) took advantage of the similar constraints on mass proxies in fossil horses to use a single proxy, m1 length, to reconstruct all the fossil genera in their study of Cope's Rule, dwarfism, and nanism in the family Equidae. Finarelli and Flynn's (2006) study of ancestral body mass in carnivorans used m1 area to reconstruct masses in fossil carnivores, demonstrating the important contribution of fossil data to accurately reconstructing the ecological history of a lineage.

Reconciling Multiple Mass Proxies

The largest extant rodent, the capybara, is dwarfed by the extinct rodents *Phoberomys* and *Josephoartigasia*. The magnitude of this size difference, however, has been a bit difficult to tie down, in part because different proxies yield different answers. *Josephoartigasia* is known primarily from craniodental remains, but the partial skeleton known for *Phoberomys pattersoni* reveal the complexity of reconstructing body sizes well outside the range encompassed by living analogs. Sánchez-Villagra and colleagues (2003) used diameters of the humerus and femur to infer mass of a specimen of *P. pattersoni*. Regressions yielded two very different answers, 436 kg for the humerus and 741 kg for the femur. The authors chose the larger size reconstructed by the femur, arguing that posture in *Phoberomys* was likely unusual relative to other caviomorphs and the femur was expected to be weight-bearing. Hopkins (2008) obtained a much smaller mass estimate (between 164 and 334 kg) from dental dimensions in a study developing tooth row dimensions as a mass proxy for rodents. Millien and Bovy (2010) reconciled the issue by comparing multiple proxies and the allometries of the proxy values relative to other skeletal measurements and by considering other information about the biology of this species. They came to the conclusion that

Phoberomys had very robust limb bones relative to the rest of its skeleton, perhaps as a result of the important postural and locomotor differences necessitated by growing rodent skeletal morphology to this extraordinary size; they concluded that the dental estimates were likely more accurate, yielding a mass estimate around 200–300 kg rather than the 741 kg proposed based on femur diameter.

Finding the Right Body Mass Regression

Freudenthal and Martín-Suárez (2013) explored some of the challenges in reconstructing body mass in extinct rodents from teeth. Using a large database of dental measurements and masses from fossil and extant rodents of a variety of lineages, the authors sought to understand which dental proxies most effectively predict mass in rodents, given that the rodent fossil record is almost exclusively dental and cranial, with few identified postcrania. The authors explored the issues around changing dental formula, comparing m1 and tooth row length as predictors of mass in rodents with and without a fourth premolar. They found that tooth row length is a strong predictor of body mass, even when calculated from the summed lengths of isolated teeth, but that there is a persistent problem with minor discrepancies in the mass reconstructed from upper and lower tooth rows. The authors explored various methods, including phylogenetic differences in the performance of the different proxies and the role of log averaging, and generated a set of regression equations that are useful for reconstructing body mass in a variety of rodent species.

Community Paleoecology

Mass proxies are commonly used in ecological diversity analyses (see Kovarovic et al. 2018) and sometimes used to reconstruct size distributions in fossil mammal communities. One technique commonly used in fossil assemblages is the cenogram (Valverde 1964), a rank-order plot of body sizes within a community sometimes inferred to reflect habitat (Legendre 1986; Travouillon and Legendre 2009) on the basis of the presence and size of breaks in slope or a gap in the size distribution. For such an application, masses are generally reconstructed using a single skeletal proxy, although the regressions (e.g., Legendre 1986) may be clade-specific. In most cases, in order to maximize the number of taxa from the fossil assemblage included, most workers have used dental proxies for mass. The cenogram method has been widely applied (Legendre 1989; Morgan et al. 1995; Croft 2001; Costeur 2004; Travouillon et al. 2009), although interpretation often requires awareness of the phylogenetic limitations of the regional species pool (e.g., Croft 2001) and may be more useful as a descriptive technique than an interpretive one (Maas and Krause 1994; Rodriguez 1999).

Future Prospects

Increasing computational power in recent years has led to an increase in the sophistication of body mass reconstructions in fossil animals. Certainly, accurate volumetric estimation would not be practical without high-powered computing (Bates et al. 2015; Brassey et al. 2015; Basu et al. 2016), although early versions of these approaches used much simpler methods with physical models (Alexander 1985). It is likely that the increasingly widespread availability of such computing power as well as the much wider access to 3D scanning will lead to more such reconstructions for species for which there are no good living analogs, either phylogenetically, morphologically, or ecologically. For the more divergent species from the fossil record such as *Sivatherium* (Basu et al. 2016), this is likely the best way to achieve a meaningful estimate of size. For most species, however, allometric equations for single proxies will remain most practical. Even these estimates can be improved by current computing capabilities, with greater access to 3D and CT scanning making it increasingly inexpensive and practical to add proxies such as articular dimensions and limb bone cortical volume to the efforts to predict mass in a greater range of species. Nonlinear regression algorithms are more sophisticated now than they were in the past, making it possible to test suggested unlogged allometric regressions (Packard 2013) to address the controversy in the log-averaging problem. It remains to be seen whether that problem will prove tractable given the substantial problems presented by heteroscedastic variance over a wide range of sizes, and whether the problem will actually translate to inaccurately-estimated masses (Kerkhoff and Enquist 2009). Finally, the growth of phylogenetic comparative methods in the last decade or two (Pennell and Harmon 2013) demands a more sophisticated analysis of the interaction between physical and physiological constraints and evolutionary history in holding these correlations together; a recent analysis of cetacean body size proxies (Pyenson and Sponberg 2011) suggests that the inclusion of phylogenetic information enhances the precision of body size estimates significantly. Given the abundance of data available and the increasingly accurate and detailed phylogenies of living species, the problem is overdue for a phylogenetic comparative treatment. A recent analysis suggests physical constraints are the primary driver of limb bone proxies (Campione and Evans 2012), but dental proxies will likely prove more complex, given the genetic constraints on dental morphology (Kangas et al. 2004; Evans et al. 2007; Kavanagh et al. 2007). A more nuanced understanding of the underlying processes can only benefit our ability to reconstruct accurate masses for extinct species.

Acknowledgments Thanks are due to Darin Croft, Scott Simpson and Denise Su for organizing the mammal paleoecology symposium that led to this volume, inspiring numerous exciting and productive conversations. I also owe thanks to Samantha Price for endless patience in teaching me phylogenetic comparative methods. Thanks to undergraduate and graduate students past and present for questioning assumptions, inspiring research, and actually getting science done. In particular, I will be ever grateful to John Orcutt for deep discussions of body size evolution in the course of his dissertation research. Finally, thanks to Edward Davis for years of discussion of body mass reconstruction, statistics of fossil samples, and regression analysis.

References

Alexander, R. M. (1985). Mechanics of posture and gait of some large dinosaurs. *Zoological Journal of the Linnean Society, 83,* 1–25.

Alroy, J. (1998). Cope's rule and the dynamics of body mass evolution in North American fossil mammals. *Science, 280,* 731–734.

Alroy, J. (2012). Simple equations for estimating body mass in mammals (and dinosaurs). *Journal of Vertebrate Paleontology, SVP Program and Abstracts Book, 2012,* 55–56.

Anderson, J. F., Hall-Martin, A., & Russell, D. A. (1985). Long-bone circumference and weight in mammals, birds and dinosaurs. *Journal of Zoology, 207,* 53–61.

Basu, C., Falkingham, P. L., & Hutchinson, J. R. (2016). The extinct, giant giraffid *Sivatherium giganteum*: skeletal reconstruction and body mass estimation. *Biology Letters, 12,* 20150940.

Bates, K. T., Falkingham, P. L., Macaulay, S., Brassey, C., & Maidment, S. C. R. (2015). Downsizing a giant: re-evaluating *Dreadnoughtus* body mass. *Biology Letters, 11,* 20150215.

Biewener, A. A. (1990). Biomechanics of mammalian terrestrial locomotion. *Science, 250,* 1097–1103.

Biknevicius, A. R. (1999). Body mass estimation in armoured mammals: cautions and encouragements for the use of parameters from the appendicular skeleton. *Journal of Zoology, 248,* 179–187.

Blackburn, T. M., & Gaston, K. J. (1994). Animal body size distributions: patterns, mechanisms and implications. *Trends in Ecology & Evolution, 9,* 471–474.

Brassey, C. A., Maidment, S. C. R., & Barrett, P. M. (2015). Body mass estimates of an exceptionally complete *Stegosaurus* (Ornithischia: Thyreophora): comparing volumetric and linear bivariate mass estimation methods. *Biology Letters, 11,* 20140984.

Brown, J. H., Gillooly, J. F., Allen, A. P., Savage, V. M., & West, G. B. (2004). Toward a metabolic theory of ecology. *Ecology, 85*, 1771–1789.

Calder, W. A. I. (1984). *Size, function, and life history*. Cambridge, MA: Harvard University Press.

Campione, N. E., & Evans, D. C. (2012). A universal scaling relationship between body mass and proximal limb bone dimensions in quadrupedal terrestrial tetrapods. *BMC Biology, 10,* 60.

Carrano, M. T., & Hutchinson, J. R. (2002). Pelvic and hindlimb musculature of *Tyrannosaurus rex* (Dinosauria: Theropoda). *Journal of Morphology, 253,* 207–228.

Christiansen, P. (2002). Mass allometry of the appendicular skeleton in terrestrial mammals. *Journal of Morphology, 251,* 195–209.

Conroy, G. C. (1987). Problems of body-weight estimation in fossil primates. *International Journal of Primatology, 8,* 115–137.

Copes, L. E., & Schwartz, G. T. (2010). The scale of it all: postcanine tooth size, the taxon-level effect, and the universality of Gould's scaling law. *Paleobiology, 36,* 188–203.

Costeur, L. (2004). Cenogram analysis of the Rudabánya mammalian community: palaeoenvironmental interpretations. *Palaeontographia Itallica, 90,* 303–307.

Croft, D. A. (2001). Cenozoic environmental change in South America as indicated by mammalian body size distributions (cenograms). *Diversity and Distributions, 7,* 271–287.

Dagosto, M., & Terranova, C. J. (1992). Estimating the body size of Eocene primates: a comparison of results from dental and postcranial variables. *International Journal of Primatology, 13,* 307–344.

Damuth, J. (1990). Problems in estimating body masses of archaic ungulates using dental measurements. In J. Damuth & B. J. MacFadden (Eds.), *Body size in mammalian paleobiology: Estimation and biological implications* (pp. 229–253). Cambridge: Cambridge University Press.

Damuth, J., & MacFadden, B. J. (Eds.). (1990). *Body size in mammalian paleobiology: Estimation and biological implications*. Cambridge: Cambridge University Press.

Delson, E., Terranova, C. J., Jungers, W. L., Sargis, E. J., Jablonski, N. G., & Dechow, P. C. (2000). Body mass in Cercopithecidae (Primates, Mammalia): estimation and scaling in extinct and extant taxa. *Anthropological Papers of the American Museum of Natural History*, *83*, 1–159.

Egi, N. (2001). Body mass estimates in extinct mammals from limb bone dimensions: the case of North American hyaenodontids. *Palaeontology, 44,* 497–528.

Eisenberg, J. F. (1981). *The mammalian radiations*. Chicago, IL: University of Chicago Press.

Evans, A. R., & Pineda-Munoz, S. (2018). Inferring mammal dietary ecology from dental morphology. In D. A. Croft, D. F. Su & S. W. Simpson (Eds.), *Methods in paleoecology: Reconstructing Cenozoic terrestrial environments and ecological communities* (pp. 37–51). Cham: Springer.

Evans, A. R., Wilson, G. P., Fortelius, M., & Jernvall, J. (2007). High-level similarity of dentitions in carnivorans and rodents. *Nature, 445,* 78–81.

Field, D. J., Lynner, C., Brown, C., & Darroch, S. A. F. (2013). Skeletal correlates for body mass estimation in modern and fossil flying birds. *PLoS ONE*, *8*, e82000.

Finarelli, J. A., & Flynn, J. J. (2006). Ancestral state reconstruction of body size in the Caniformia (Carnivora, Mammalia): the effects of incorporating data from the fossil record. *Systematic Biology, 55,* 301–313.

Fortelius, M. (1985). Ungulate cheek teeth: developmental, functional, and evolutionary interrelations. *Acta Zoologica Fennica, 180,* 1–76.

Fortelius, M. (1990). Problems with using fossil teeth to estimate body sizes of extinct mammals. In J. Damuth & B. J. Macfadden (Eds.), *Body size in mammalian paleobiology: Estimation and biological implications* (pp. 207–228). Cambridge: Cambridge University Press.

Freudenthal, M., & Martín-Suárez, E. (2013). Estimating body mass of fossil rodents. *Scripta Geologica, 14,* 1–130.

Freudenthal, M., & Martín-Suárez, E. (2015). Estimating head and body length in fossil rodents. *Scripta Geologica, 149,* 1–158.

Gillooly, J. F., Allen, A. P., West, G. B., & Brown, J. H. (2005). The rate of DNA evolution: effects of body size and temperature on the molecular clock. *Proceedings of the National Academy of Sciences, USA, 102,* 140–145.

Gingerich, P. D. (1974). Size variability of the teeth in living mammals and the diagnosis of closely related sympatric fossil species. *Journal of Paleontology, 48,* 895–903.

Gingerich, P. D. (1990). Prediction of body mass in mammalian species from long bone lengths and diameters. *Contributions from the Museum of Paleontology, the University of Michigan, 28,* 79–92.

Gingerich, P. D. (2000). Arithmetic or geometric normality of biological variation: an empirical test of theory. *Journal of Theoretical Biology, 204,* 201–221.

Gingerich, P. D., Smith, B. H., & Rosenberg, K. (1982). Allometric scaling in the dentition of primates and prediction of body weight from tooth size in fossils. *American Journal of Physical Anthropology, 58,* 81–100.

Glazier, D. S. (2013). Log-transformation is useful for examining proportional relationships in allometric scaling. *Journal of Theoretical Biology, 334,* 200–203.

Gordon, C. L. (2003). A first look at estimating body size in dentally conservative marsupials. *Journal of Mammalian Evolution, 10,* 1–21.

Gould, G. C., & MacFadden, B. J. (2004). Gigantism, dwarfism, and Cope's Rule: "Nothing in evolution makes sense without a phylogeny". *Bulletin of the American Museum of Natural History, 285,* 219–237.

Gould, S. J. (1975). On the scaling of tooth size in mammals. *American Zoologist, 15,* 351–362.

Hopkins, S. S. B. (2008). Reassessing the mass of exceptionally large rodents using toothrow length and area as proxies for body mass. *Journal of Mammalogy, 89,* 232–243.

Hutchinson, J. R., & Garcia, M. (2002). *Tyrannosaurus* was not a fast runner. *Nature, 415,* 1018–1021.

Huxley, J. S. (1932). *Problems of relative growth*. London: Methuen & Co., Ltd.

Iskjaer, C., Slade, N. A., Childs, J. E., Glass, G. E., & Korch, G. W. (1989). Body mass as a measure of body size in small mammals. *Journal of Mammalogy, 70,* 662–667.

Janis, C. M. (1990). Correlation of cranial and dental variables with body size in ungulates and macropodoids. In J. Damuth & B. J. MacFadden (Eds.), *Body size in mammalian paleobiology: Estimation and biological implications* (pp. 255–300). Cambridge: Cambridge University Press.

Kangas, A. T., Evans, A. R., Thesleff, I., & Jernvall, J. (2004). Nonindependence of mammalian dental characters. *Nature, 432,* 211–214.

Kaufman, J. A., & Smith, R. J. (2002). Statistical issues in the prediction of body mass for Pleistocene canids. *Lethaia, 35,* 32–34.

Kavanagh, K. D., Evans, A. R., & Jernvall, J. (2007). Predicting evolutionary patterns of mammalian teeth from development. *Nature, 449,* 427–432.

Kay, R. F., & Ungar, P. S. (1997). Dental evidence for diet in some Miocene catarrhines with comments on the effects of phylogeny on the interpretation of adaptation. In D. R. Begun, C. V. Ward & M. D. Rose (Eds.), *Function, phylogeny, and fossils: Miocene hominoid evolution and adaptations* (pp. 131–151). Dordrecht: Springer.

Kerkhoff, A. J., & Enquist, B. J. (2009). Multiplicative by nature: why logarithmic transformation is necessary in allometry. *Journal of Theoretical Biology, 257,* 519–521.

Kovarovic, K., Su, D. F., & Lintulaakso, K. (2018). Mammal community structure analysis. In D. A. Croft, D. F. Su & S. W. Simpson (Eds.), *Methods in paleoecology: Reconstructing Cenozoic terrestrial environments and ecological communities* (pp. 349–370). Cham: Springer.

LaBarbera, M. (1989). Analyzing body size as a factor in ecology and evolution. *Annual Review of Ecology and Systematics, 20,* 97–117.

Legendre, S. (1986). Analysis of mammalian communities from the late Eocene and Oligocene of southern France. *Palaeovertebrata, 16,* 191–212.

Legendre, S. (1989). Les communautés de mammifères du Paléogène (Eocène supérieur et Oligocène) d'Europe occidentale : Structures, milieux et évolution. *Münchner Geowissenschaftliche Abhandlungen A, 16,* 1–110.

Lindsay, E. H. (1988). Cricetid rodents from Siwalik deposits near Chinji Village. Part 1: Megacricetodontinae. Myocricetodontinae and Dendromurinae. *Palaeovertebrata, 18,* 95–154.

Lindstedt, S. L., & Boyce, M. S. (1985). Seasonality, fasting endurance, and body size in mammals. *The American Naturalist, 125,* 873–878.

Liow, L. H., Fortelius, M., Bingham, E., Lintulaakso, K., Mannila, H., Flynn, L., et al. (2008). Higher origination and extinction rates in larger mammals. *Proceedings of the National Academy of Sciences, USA, 105,* 6097–6102.

Liow, L. H., Fortelius, M., Lintulaakso, K., Mannila, H., & Stenseth, N. C. (2009). Lower extinction risk in sleep-or-hide mammals. *The American Naturalist, 173,* 264–272.

Lockyer, C. (1976). Body weights of some species of large whales. *Journal du Conseil / Conseil Permanent International pour l'Exploration de la Mer, 36,* 259–273.

Maas, M. C., & Krause, D. W. (1994). Mammalian turnover and community structure in the Paleocene of North America. *Historical Biology, 8,* 91–128.

Martin, A. P., & Palumbi, S. R. (1993). Body size, metabolic rate, generation time, and the molecular clock. *Proceedings of the National Academy of Sciences, USA, 90,* 4087–4091.

Martin, R. A. (1980). Body mass and basal metabolism of extinct mammals. *Comparative Biochemistry and Physiology – Part A. Physiology, 66,* 307–314.

Martin, R. A. (1990). Estimating body mass and correlated variables in extinct mammals: travels in the fourth dimension. In J. Damuth & B. J. MacFadden (Eds.), *Body size in mammalian paleobiology: Estimation and biological implications* (pp. 49–68). Cambridge: Cambridge University Press.

McMahon, T., & Bonner, J. T. (1983). *On size and life*. New York: Scientific American Books – W. H. Freeman and Co.

McNab, B. K. (1963). Bioenergetics and the determination of home range size. *The American Naturalist, 97,* 133–140.

McNab, B. K. (1988). Complications inherent in scaling the basal rate of metabolism in mammals. *The Quarterly Review of Biology, 63,* 25–54.

Mendoza, M., Janis, C. M., & Palmqvist, P. (2006). Estimating the body mass of extinct ungulates: a study on the use of multiple regression. *Journal of Zoology, 270,* 90–101.

Millien, V., & Bovy, H. (2010). When teeth and bones disagree: body mass estimation of a giant extinct rodent. *Journal of Mammalogy, 91,* 11–18.

Morgan, M. E., Badgley, C., Gunnell, G. F., Gingerich, P. D., Kappelman, J. W., & Maas, M. C. (1995). Comparative paleoecology of Paleogene and Neogene mammalian faunas: body-size structure. *Palaeogeography, Palaeoclimatology, Palaeoecology, 115,* 287–315.

Myers, T. J. (2001). Prediction of marsupial body mass. *Australian Journal of Zoology, 49,* 99–118.

Packard, G. C. (2009). On the use of logarithmic transformations in allometric analyses. *Journal of Theoretical Biology, 257,* 515–518.

Packard, G. C. (2013). Is logarithmic transformation necessary in allometry? *Biological Journal of the Linnean Society, 109,* 476–486.

Pennell, M. W., & Harmon, L. J. (2013). An integrative view of phylogenetic comparative methods: connections to population genetics, community ecology, and paleobiology. *Annals of the New York Academy of Sciences, 1289,* 90–105.

Peters, R. H. (1983). *The ecological implications of body size*. Cambridge: Cambridge University Press.

Price, S. A., & Hopkins, S. S. B. (2015). The macroevolutionary relationship between diet and body mass across mammals. *Biological Journal of the Linnean Society, 115,* 173–184.

Pyenson, N. D., & Sponberg, S. N. (2011). Reconstructing body size in extinct crown Cetacea (Neoceti) using allometry, phylogenetic methods and tests from the fossil record. *Journal of Mammalian Evolution, 18,* 269–288.

Rafferty, K. L., Walker, A., Ruff, C. B., Rose, M. D., & Andrews, P. J. (1995). Postcranial estimates of body weight in *Proconsul*, with a note on a distal tibia of *P. major* from Napak, Uganda. *American Journal of Physical Anthropology*, *97*, 391–402.

Reynolds, P. S. (2002). How big is a giant? The importance of method in estimating body size of extinct mammals. *Journal of Mammalogy, 83,* 321–332.

Ricker, W. E. (1973). Linear regression in fishery research. *Journal of the Fisheries Research Board of Canada, 30,* 409–434.

Ricker, W. E. (1984). Computation and uses of central trend lines. *Canadian Journal of Zoology, 62,* 1897–1905.

Rinderknecht, A., & Blanco, R. E. (2008). The largest fossil rodent. *Proceedings of the Royal Society B: Biological Sciences, 275,* 923–928.

Rodriguez, J. (1999). Use of cenograms in mammalian palaeoecology. A critical review. *Lethaia, 32,* 331–347.

Roth, V. L. (1990). Insular dwarf elephants: a case study in body mass estimation and ecological inference. In J. Damuth & B. J. Macfadden (Eds.), *Body size in mammalian paleobiology: Estimation and biological implications* (pp. 151–180). Cambridge: Cambridge University Press.

Roth, V. L. (1992). Inferences from allometry and fossils: dwarfing of elephants on islands. *Oxford Surveys in Evolutionary Biology, 8,* 259–288.

Ruff, C. (1988). Hindlimb articular surface allometry in Hominoidea and *Macaca*, with comparisons to diaphyseal scaling. *Journal of Human Evolution, 17,* 687–714.

Ruff, C. B. (1990). Body mass and hindlimb bone cross-sectional and articular dimensions in anthropoid primates. In J. Damuth & B. J. Macfadden (Eds.), *Body size in mammalian paleobiology: Estimation and biological implications* (pp. 119–150). Cambridge: Cambridge University Press.

Ruff, C. B. (2002). Long bone articular and diaphyseal structure in old world monkeys and apes. I: locomotor effects. *American Journal of Physical Anthropology, 119,* 305–342.

Ruff, C. B. (2003). Long bone articular and diaphyseal structure in Old World monkeys and apes. II: estimation of body mass. *American Journal of Physical Anthropology, 120,* 16–37.

Ruff, C. B., Scott, W. W., & Liu, A. Y.-C. (1991). Articular and diaphyseal remodeling of the proximal femur with changes in body mass in adults. *American Journal of Physical Anthropology, 86,* 397–413.

Sánchez-Villagra, M. R., Aguilera, O., & Horovitz, I. (2003). The anatomy of the world's largest extinct rodent. *Science, 301,* 1708–1710.

Schmidt-Nielsen, K. (1984). *Scaling: Why is animal size so important?* Cambridge: Cambridge University Press.

Schulte-Hostedde, A. I., Zinner, B., Millar, J. S., & Hickling, G. J. (2005). Restitution of mass-size residuals: validating body condition indices. *Ecology, 86,* 155–163.

Scott, J. E. (2011). Folivory, frugivory, and postcanine size in the cercopithecoidea revisited. *American Journal of Physical Anthropology, 146,* 20–27.

Scott, K. M. (1983). Prediction of body weight of fossil Artiodactyla. *Zoological Journal of the Linnean Society, 77,* 199–215.

Scott, K. M. (1990). Postcranial dimensions of ungulates as predictors of body mass. In J. Damuth & B. J. Macfadden (Eds.), *Body size in mammalian paleobiology: Estimation and biological implications* (pp. 301–336). Cambridge: Cambridge University Press.

Sibly, R. M., & Brown, J. H. (2007). Effects of body size and lifestyle on evolution of mammal life histories. *Proceedings of the National Academy of Sciences, USA, 104,* 17707–17712.

Smith, F. A., Boyer, A. G., Brown, J. H., Costa, D. P., Dayan, T., Ernest, S. K. M., et al. (2010). The evolution of maximum body size of terrestrial mammals. *Science, 330,* 1216–1219.

Smith, R. J. (1993). Logarithmic transformation bias in allometry. *American Journal of Physical Anthropology, 90,* 215–228.

Smith, R. J. (2002). Estimation of body mass in paleontology. *Journal of Human Evolution, 43,* 271–287.

Toigo, C., Gaillard, J. M., Van Laere, G., Hewison, M., & Morellet, N. (2006). How does environmental variation influence body mass, body size, and body condition? Roe deer as a case study. *Ecography, 29,* 301–308.

Travouillon, K. J., & Legendre, S. (2009). Using cenograms to investigate gaps in mammalian body mass distributions in Australian mammals. *Palaeogeography, Palaeoclimatology, Palaeoecology, 272,* 69–84.

Travouillon, K. J., Legendre, S., Archer, M., & Hand, S. J. (2009). Palaeoecological analyses of Riversleigh's Oligo-Miocene sites: implications for Oligo-Miocene climate change in Australia. *Palaeogeography, Palaeoclimatology, Palaeoecology, 276,* 24–37.

Valverde, J. A. (1964). Remarques sur la structure et l'évolution des communautés de vertébrés terrestres. *La Terre et La Vie, 111,* 121–154.

Van Valkenburgh, B. (1990). Skeletal and dental predictors of body mass in predators. In J. Damuth & B. J. Macfadden (Eds.), *Body size in mammalian paleobiology: Estimation and biological implications* (pp. 181–206). Cambridge: Cambridge University Press.

Van Valkenburgh, B., Wang, X., & Damuth, J. (2004). Cope's Rule, hypercarnivory, and extinction in North American canids. *Science, 306,* 101–104.

Vinyard, C. J., & Hanna, J. (2005). Molar scaling in strepsirrhine primates. *Journal of Human Evolution, 49,* 241–269.

West, G. B., Brown, J. H., & Enquist, B. J. (1997). A general model for the origin of allometric scaling laws in biology. *Science, 276,* 122–126.

White, C. R., Blackburn, T. M., & Seymour, R. S. (2009). Phylogenetically informed analysis of the allometry of mammalian basal metabolic rate supports neither geometric nor quarter-power scaling. *Evolution, 63,* 2658–2667.

Whittaker, R. J. (1999). Scaling, energetics, and diversity. *Nature, 401,* 865–866.

Wu, C. F. J. (1986). Jackknife, bootstrap and other resampling methods in regression analysis. *The Annals of Statistics, 14,* 1261–1295.

Xiao, X., White, E., Hooten, M., & Durham, S. (2011). On the use of log-transformation vs. nonlinear regression for analyzing biological power-laws. *Ecology, 92,* 1887–1894.

Zar, J. H. (2010). *Biostatistical analysis (5th ed)*. Englewood Cliffs: Prentice-Hall.