Part I

Background

1 Introduction

There are two. . . schools of biologists, one which emphasizes the average and general and the other the individual and particular.

Brody 1945

My interest in *scaling* in mammals was first sparked when, as a fledgling graduate student adapting fitfully to a human-centered and largely qualitative medical school environment, I read a paper by Günther and Guerra kindly passed on to me by my thesis advisor, Alan Burton. This paper showed (as was already well known to others) that there are quantitative regularities in the construction of mammals of varied sizes not apparent from studies limited to human anatomy and physiology. Little did I then realize that scaling would subsequently come to occupy such a large and beguiling part of my adult vocation.

The term scaling is concerned with how the properties and proportions of a given system vary (or fail to vary) with system size or with time. Discussions of so-called *spatial* and *temporal* scaling are common to many disciplines. It should be said that scaling is notable for its breadth of applications more than its depth. Scaling studies, in the first instance, tend to be descriptive rather than explanatory. The term "system" is open to different interpretations in fields as diverse as architecture, astronomy, climate change, ecology, economics, engineering, landscape design, paleontology, physics, and physiology. An important assumption in scaling studies, usually implicit rather than explicit, is that two or more examples of a given system differing in size (or time interval) are qualitatively similar with respect to the properties under study. That is to say, variation in these properties is strictly quantitative. It is generally accepted that regular scaling with size in any system always breaks down at sufficiently small or large sizes because of *scale effects* (see Chapter 3).

Scaling in biology has been defined as the study of body *size* and its consequences. It hardly needs saying that body size affects virtually every aspect of mammalian life. Adult mammals vary in size by a factor of 100 million (from bumblebee bat to blue whale; *Craseonycteris thonglongyai* to *Balaenoptera musculus*). Apart from fish, no other vertebrate group, living or extinct, spans a comparable size range (Chapter 2). Notwithstanding this remarkable 100-million-fold size range, all mammals broadly share the same cell types, the same tissues and organs, and the same overall body plan. Nevertheless, no one seriously believes that scale-up by a factor of a hundred million in a complex system can be achieved without a diverse suite of internal and external

changes, both structural and functional. These adjustments to increased body size are brought into being by natural selection.

Hence, scaling in mammals may be defined as the quantitative study of those adjustments (or self-adaptations) as well as non-adjustments made in response to increased (or decreased) body size, mediated by natural selection (rather than size per se). Scaling studies stand to reveal the many respects in which adult mammals generally are similar as well as (in perhaps fewer but still important respects) dissimilar.

It follows that a major aim of mammalian scaling studies is to determine, and ultimately to understand, the nature of these – frequently very regular – internal and external changes made as a function of changing body size in mammals generally. These studies may be developmental (ontogenetic) in a single species, adult interspecific (across multiple species) or, as is often the case in this study, both. We will see that these internal changes involve, for example, at increasing body sizes, a reallocation of some biological resources away from energetically expensive organs and tissues (e.g., brain, kidney, liver) to less expensive ones (e.g., bone and fat). Thus, scaling in mammals is concerned in part with what may be termed "bio-economics."

The long-term goal of scaling studies in mammals is to describe the adjustments and non-adjustments to changes in body size, to identify the underlying mechanisms of adjustment, to estimate when, in the course of evolution, the adjustments were made and to determine why, insofar as that may prove possible, certain adjustments were made but not others. Such explorations stand to appreciably deepen our understanding of mammalian "design."

Mammalian scaling studies usually involve two variables, here denoted by x and y. In the first instance, we aim to answer empirically the question of how y scales (varies) with x. Commonly, the x-variable is a measure of body or organ size. The y-variable may be either structural or functional. Examples of *structural* y-variables include the number of components, body length, body surface area, and organ volume. These y-variables have dimensions of zero (number), one (length), two (area), and three (volume), respectively. Instances of *functional* y-variables include blood pressure, body temperature, and heart rate. Generally speaking, functional y-variables are dimensionally more disparate than structural variables.

In practice we plot the logarithm of a y-variable against the logarithm of an x-variable. The answer to the initial question of how y varies with x is given by the *slope* (b) of the straight line expressing y as a function of x. In the expression for a power function, given by $y = ax^b$, the constant a is numerically equal to the antilog of the intercept (i.e., where $x = 1$) and the exponent b to the slope of the straight line, assuming that x and y are given in log–log coordinates. The slope may be viewed (see Chapter 23) as an average – or perhaps better as a median – and as a generalization (or simplification). Scaling studies inherently involve quantitative generalizations (see above quote from Brody).

This work on scaling (see Contents) is divided into five Parts. Proceeding in reverse order, Part V (Chapter 25) takes a broader view of scaling, by examining the potential application of scaling concepts to currently pressing global problems – climate change, soil erosion, and water shortages. Methodological aspects of scaling studies in

mammals are discussed in Part IV (Chapters 23, 24). Methodology is important, because numbers are no better than the methods used to acquire them.

A concise summary of many – but not all – of the results derived in Part II is provided in Part III (Chapters 19 to 22). The possible influences of dimensions and invariance on empirical slopes are discussed. Part II itself (Chapters 6 to 18), the core of this work, brings together the results of an extensive empirical analysis of scaling in mammals, involving approximately 100 different *y*-variables. The organization of the work is intended to facilitate finding those results in which the reader may be most interested.

Part II is organized mainly on a *systemic* basis. Recall that for more than a century mammals have been analyzed mainly in terms of subsystems. This is reflected, for example, in the specialities of cardiovascular, respiratory, and urinary anatomy and physiology. Whole body functions such as body temperature and energy metabolism are discussed in Chapter 17. Part II closes (Chapter 18) with a discussion of lethal limits. The way in which lethal limits scale may be useful for understanding how exacting are homeostatic constraints when placed under external stress.

Part I (Chapters 1 to 5) presents background material intended to facilitate an understanding of the (partly novel) presentations given in Parts II and III. Keep in mind that the present study does not attempt to make either ecological or phylogenetic inferences. On the contrary, the aim here is to describe, and eventually understand, structural and functional scaling in mammals *generally*, independent, so far as possible, of ecological and phylogenetic considerations. For example, water molecules, which make up about two-thirds of an adult mammal, are all the same in size and shape. Increase in adult body size across species entails an increase in the *number* of water molecules, all identical to one another. This well-ordered increase, like numerous others, appears to be essentially independent, on average, of either ecology or phylogeny.

Adult mammals vary both *taxonomically* and in *body size*. It is helpful to have a quantitative measure of diversity that tells us how nearly a given dataset comes to representing mammals generally or whether one dataset is more representative of mammals than another. We first consider *taxonomic* variation. Let us denote the number of species in a given sample by N_{sp} and the number of orders by N_{ord}. We take taxonomic variation into account by comparing the number of species and orders present in a dataset with the number of species and orders of mammals generally (see Chapter 2).

A measure of *size variation* is provided by pWR, calculated as the logarithm of the ratio of the largest body mass in a sample to the smallest. We know (see above) that pWR for mammals generally is about eight (pWR = $\log(10^8) = 8$). In Part I, Chapter 2, the values of N_{sp}, N_{ord}, and pWR are combined to provide a single *diversity index* (DI); this index is used throughout the work. These four parameters (N_{ord}, N_{sp}, pWR, DI) are collectively referred to as evaluative criteria.

It will be found that *structural dimensions* (for number, length, area, and volume or mass) are of considerable relevance to scaling in adult mammals. Much of Part II is concerned with how these dimensions scale. *Invariance* is also of particular importance in scaling studies. Invariance simply means that a given property of a system remains constant (invariant) under a specific transformation. In a broad sense invariance speaks

to the issue of "constant design." We may define two types of biological invariance: *absolute* and *relative* (Chapter 3). Absolute invariance means that component attributes such as number, diameter, area, and volume or mass show little if any dependence on adult body size. In Part II we encounter absolute invariance for the mean size and shape of most types of small and large molecules (e.g., glucose, hemoglobin, water) and for mean size and shape of some cell types (e.g., red blood cells). Mean body temperature and mean blood pressure and other homeostatic variables may exhibit absolute or near-absolute invariance. It will be argued that absolute invariance, both structural and functional, is relatively common in mammalian scaling.

It is also useful in considerations of mammalian *structural* scaling to define *relative invariance* (Chapter 3). This term means that a structural parameter (e.g., mass, volume or component number) increases in direct proportion to body size. For example, blood volume, the mass of the lungs, and the mass of skeletal muscle in adult mammals scale in reasonable accord, on average, with relative invariance. The same is likely to be true of the total number of water molecules or red blood cells in the adult mammal. Absolute and relative invariance are of considerable importance in mammalian scaling studies (see Chapter 3). Observe that some forms of absolute invariance are physically determined, while other forms are *elective* (subject to natural selection). For example, the size of water molecules is physically determined, whereas the size of hemoglobin molecules is probably elective. In all probability, all forms of relative invariance are elective.

The scaling parameter of most interest is the *slope* (b) of the best-fit line expressing y as a function of x, both in log–log coordinates. If we are told that a given best-fit line has a slope of 2/3, we want to know the implications of that number for the scaling of the y-variable. Strictly speaking, it is the meaning of numbers – not the numbers themselves – that is of scientific importance. Perhaps the simplest approach to understanding the implication of a slope is to compute the *factor* by which the y-variable changes for a given *factor* of change in the x-variable (see Chapter 3). If, for example, the x-variable doubles, what is the factor by which the y-variable changes?

Commonly we have two or more independent datasets specifying how a given y-variable scales with the x-variable. If, as is usually the case, two specific datasets give rise to best-fit lines with somewhat different slopes, we want to know the scaling implications of this *difference* in slope. We are interested in the scaling implications of a difference in slope for scaling in mammals generally. That is, we elect to assess the impact of a difference in slope on the *hypothesis* that the best-fit lines apply over the *whole* mammalian size range (pWR = 8) (see above). We provide three different assessments of the discrepancy between the end-points of two lines with different slopes assumed to have a common intercept (Chapter 3). The first assessment employs *percent difference*; the second is expressed in terms of *number of doublings*; the third is based on the *factor* by which one end-point is larger than the other. These three assessments are inter-related.

Authors who provide scaling data (i.e., measures of some y-variable with corresponding measures of body size) sometimes state that the body size measurements were made on *adult* mammals. Often this consideration is ignored. Further, when animals are said

to be adult it is uncommon to find a statement of the *criterion* used to make this judgement. A reasonable assumption is that adulthood is regularly equated with sexual maturity. We know that for numerous species, including our own, sexual maturity is reached before physical maturity. The danger for scaling studies is that a systematic bias in body mass may be introduced if small species are truly adults and large species are not, or vice versa.

In order to reduce the magnitude of this problem, a computer-based body weight (or mass) table was constructed (largely completed by 1995). Frequently, several estimates of body weight (or mass) from different works were available for a given species. In those cases, representative body weights were selected that were closest to the mean of the several estimates. The database constructed in this way is here referred to, for convenience, as the *standardized body weight table*, denoted by *SBT* (see Chapter 4 and Appendix B). The standardized body weight table contains body weight (or mass) data for some 1,700 species of mammals.

The great majority of datasets analyzed in Part II were *screened* using the standardized body weight table and the criteria given in Appendix B. This means that when a dataset from a particular author was screened the resultant dataset was usually somewhat smaller than the one originally reported by the author. On the other hand, my aggregated datasets, as a rule drawing on screened data from multiple authors, are usually larger than those previously reported as of the time of my analysis.

When one derives a *best-fit line*, usually by least squares (LSQ), for a given scaling dataset, we need *evaluative* criteria by which to judge how well a best-fit line describes the underlying dataset. I have used three different criteria: *mean percent deviation* (MPD), the *coefficient of determination* (r^2), and the *99% confidence limits* (CL) (in most cases) on the slope (see Chapter 5). The MPD provides a *non-dimensional* measure of the scatter around the best-fit line. The coefficient of determination tells how much of the variation in the y-variable is attributable to variation in the x-variable. The 99% CL provide a measure of the confidence we can place in the slope. Thus the great majority of best-fit lines reported in Part II are characterized by a total of seven *evaluative* parameters (N_{ord}, N_{sp}, MPD, pWR, DI, r^2, and the 99% CL on the slope). For a brief review of these evaluative criteria and other matters see Chapters 3–5.

The above seven evaluative criteria do not bear on the question of how *sensitive* a slope derived by LSQ is to sample size and to the range in body size. To examine this question I carried out the following experiments. First, I looked at 24 datasets (18 structural and 6 functional) based on 100 (or more) entries (species). For each of these 24 datasets (not including the separate case of body length – see Chapter 6) I constructed an *end-sample*, consisting of just those records for the ten smallest and the ten largest species by body mass (i.e., 20 records in all). Then a *mid*-sample was derived by deleting from the full dataset (FDS) the ten smallest and ten largest species. Thus, a mid-sample comprised 80 or more records. A LSQ analysis was carried out for each end-sample, mid-sample, and FDS. For the possibly unexpected results of these experiments see Chapters 6 and 21.

The analytical work underlying this study proceeded in two overlapping phases. The first phase, spread over several decades, was *bibliographic*. All the relevant references

were entered into a computer-based bibliographic program. Beyond the many papers widely cited in the mammalian scaling literature, many other useful papers (especially in the biomedical literature) were found that have rarely been cited. Bibliographic research on a given topic (say the circulatory system) was curtailed when diminishing returns set in. The aggregate bibliographic database contains more than 10,000 references. Many, but not all, of these references are directly pertinent to scaling in mammals.

The second phase of the work consisted of *numerical analysis*. Raw scaling data on a given topic, normally from multiple different sources, were entered into a spreadsheet program. In nearly all cases the "raw" data were screened for body weight or mass and by other criteria (see Appendix B). After screening, a best-fit line was computed using LSQ. Once the above two phases had been completed for a given topic, further work was generally not undertaken. As a consequence, the coverage of some topics is more up-to-date than for others.

Note that each dataset consists of some number of records (or rows in a spreadsheet). For most datasets the number of records is the same as the number of species. But two datasets – here termed "instances" – for the same *y*-variable (e.g., body length) will usually differ in the number of records and in the spectrum of species. Nonetheless, two datasets for the same *y*-variable will commonly contain some records for the same species. Thus not all the records in two (or more) datasets will be unique. Part II is based on the analysis of roughly 16,000 records in all, for nearly 100 different *y*-variables. Part III is based on roughly 7,500 records (all unique) for 72 *y*-variables (52 structural and 20 functional), all in tabular form.

This is the first extensive scaling study devoted almost exclusively to mammals. It is the first to assemble a large number of databases and to *analyze* them uniformly, using LSQ analysis and seven different *evaluative criteria* (see above). Likewise, it is the first to take a systemic approach to scaling for eight different subsystems, running from the circulatory to the urinary (see Contents). It is the first study to compare and contrast physical and biological scaling in some detail. Also new is the extensive comparison between *ontogenetic* (chiefly human) scaling and *adult interspecific* scaling.

In addition, this work is apparently the first to compare adult humans with other adult mammals across 28 structural variables (see Chapter 22). For an extended discussion of body composition (elements to organs) see Chapter 8. Lethal limits are reviewed at greater length than previously in Chapter 18. Finally, there is a wider discussion of scaling methodology (see Chapters 3, 23, 24). In Chapter 23 it is shown that methods of analysis other than least squares may give similar values for the slope.

Whether the reader coming to this book is experienced in scaling studies, or a relative novice, it may prove useful to read through the chapters of Part I first before proceeding to other parts of the work. The primary purpose of Part I is to provide essential information needed to quickly and easily absorb the tabular and figurative material presented in the later chapters. (For the reader's convenience, this information is itemized in Chapter 5.) The subsequent chapters (6 through 25) are largely independent of one another. They may be read in any order. The price of this independence is a certain amount of repetition, designed to reduce (but not eliminate) the need for

cross-checking across chapters. Chapters 6 through 24 each provide a summary statement at the end of the chapter.

With three exceptions (Chapters 6, 7, 18), each chapter in Part II also provides a summary figure showing the distribution of slopes encountered in the given system expressed in terms of rank order. Those distributions that are especially distinctive are compared in Chapter 20. In the long run, the pattern of slopes for each system is likely to prove as interesting and important as the narrower study of individual slopes. A partial interpretation of these distributions of slopes is given throughout the text in terms of absolute and relative invariance as well as in terms of those dimensions associated with number, diameter, area, and volume (e.g., see Chapter 12, Figure 12.3). A significant challenge for future work related to scaling in mammals will be to discover explanations for these rather varied distributions of slope.

An engaged reader who embarks on this excursion through the varied terrain of scaling in mammals should emerge with a detailed picture of a selection of the manifold adjustments mammals make to increased body size. Just as important as structural and functional *variance* are the many ways in which mammals exhibit *invariance*, on average, despite enormous changes in body size. Like newly-weds, variance and invariance go hand in hand.

2 The mammals

Mammals are the most diversified of all creatures living on earth today.
Boitani and Bartoli 1982

The time scale

Our universe is 13.8 billion years old; planet earth coalesced 4.5 billion years ago. Fossilized stromatolites, the products of colonial cyanobacteria, are dated to 3.4 billion years ago. (Contemporary stromatolites may be seen in Australia and elsewhere.) They are the earliest widely accepted evidence of life on earth. Eukaryotes (nucleated cells) first appeared 1.5 billion years ago. Vertebrate fossils recently found in China date to 450 million years ago (mya). The earliest mammal-like reptiles, the cynodonts (Order Therapsida, Class Synapsida), surfaced in the fossil record of the Late Permian and early Triassic periods (245 mya), slightly after the dinosaurs (in geological terms). Until near the end of the Cretaceous period (65 mya) mammals were shrew- to rat-sized creatures, slinking about in the dark to evade the dominating dinosaurs. Within ten million years – a wink in the timespan of the universe – following the dinosaur extinction a significant radiation of the mammals took place, giving rise to all the contemporary taxonomic orders, with some individual species of large size [1].

The primary purpose of this chapter is to introduce some qualitative but mainly quantitative measures of mammalian diversity, both size- and taxonomy-related. These measures involve a somewhat different way of looking at mammals than has been customary in scaling studies.

Mammals as a divergent vertebrate

Structural features believed to be unique to mammals are recapped in Table 2.1; of the ten structures listed, seven are considered to be universal. In principle any one of these structures might be regarded as *the* defining characteristic of mammals (see above quote from Boitani and Bartoli). But mammary glands are the most important of these structures from an evolutionary standpoint; lactation has profound implications for mammalian social behavior [2].