Supply–demand balance and metabolic scaling

Jayanth R. Banavar*, John Damuth†‡, Amos Maritan§¶, and Andrea Rinaldo‖

*Department of Physics, 104 Davey Laboratory, Pennsylvania State University, University Park, PA 16802; †Department of Ecology, Evolution, and Marine Biology, University of California, Santa Barbara, CA 93106; §International School for Advanced Studies (S.I.S.S.A.), Via Beirut 2-4, 34014 Trieste, Italy; ¶Istituto Nazionale di Fisica della Materia and the Abdus Salam International Center for Theoretical Physics, 34100 Trieste, Italy; and ‖Dipartimento di Ingegneria Idraulica, Marittima e Geotecnica, Universitá di Padova, I-35131 Padova, Italy

It is widely accepted that metabolic rates scale across species approximately as the 3/4 power of mass in most if not all groups of organisms. Metabolic demand per unit mass thus decreases as body mass increases. Metabolic rates reflect both the ability of the organism's transport system to deliver metabolites to the tissues and the rate at which the tissues use them. We show that the ubiquitous 3/4 power law for interspecific metabolic scaling arises from simple, general geometric properties of transportation networks constrained to function in biological organisms. The 3/4 exponent and other observed scaling relationships follow when mass-specific metabolic demands match the changing delivery capacities of the network at different body sizes. Deviation from the 3/4 exponent suggests either inefficiency or compensating physiological mechanisms. Our conclusions are based on general arguments incorporating the minimum of biological detail and should therefore apply to the widest range of organisms.

Living organisms span an impressive range in body mass, varying over some 21 orders in magnitude (1, 2) from mycoplasma (10^{-13} g) to the blue whale (10^8 g). Biologists have described a large number of relationships that link body size to organ sizes, rates of physiological processes, and biological cycle times (3–5). These relationships usually take the form of power laws, allometric scaling relationships of the form $Y = aM^b$, where M is body mass, Y is the biological property of interest, and a and b are constants specific to the relationship.

Strikingly, the exponents b of a wide variety of biological properties appear to be obtainable from the number 4, even though the systems are three dimensional (3, 4, 6). The reason that so many properties exhibit this unexpected scaling is presumably their relationship to, or dependence on, the rates of metabolic processes. Whole-body metabolic rates themselves generally scale with an exponent b approximately equal to 3/4, and this well-known phenomenon is widespread among unicellular and multicellular animals and plants (3, 4, 6). Although individual datasets comprising ecologically and physiologically heterogeneous species often show some variation (3, 7, 8), the 3/4 exponent is widely accepted as a near-universal general description of the interspecific scaling of metabolism (1–6, 9, 10). That scaling exponents related to the number 4 are so pervasive in nature suggests that the underlying explanation ought to be general and simple, and not dependent on the details of the metabolic machinery of particular types of organisms.

There has been a flurry of recent activity (11–14) aimed at providing possible explanations for the exponents observed in the scaling of metabolism. Dodds *et al.* (8) have carried out a reexamination of recent theories and conclude that none of them is convincing. Our goal is to attack this open theoretical problem and present a general explanation for observed scaling relationships. Our work builds on a theorem pertaining to the scaling of general networks (14). In this paper, we pay careful attention to the specific constraints imposed on such networks when they are functioning components of the metabolic systems of biological organisms. We describe precise conditions under which 3/4 metabolic scaling would be expected, without reference to non-Euclidean geometry, "fractal-like" branching networks, or other specific (and rather exotic) features that have been claimed

(6, 11–13) to be required for organisms to exhibit 3/4 scaling. Our theory is based on familiar geometric principles and physical constraints that apply to living and nonliving systems alike.

Metabolism in most organisms entails the transportation of metabolites to the various regions of the body from an internal source. For example, in vertebrates and many other organisms, a heart pumps blood, which carries oxygen to the tissues of the body, where the oxygen is used in the biochemical processes of respiration. For convenience, in the rest of the paper we will allude to "blood" as the fluid carrying the metabolites, but the same considerations would apply for any fluids carrying any kind of metabolites that potentially limit the rate of metabolism. Note also that we assume that metabolite concentration in the blood is independent of body mass, as is the case among vertebrates (4). Likewise, we will assume that the energy capacity "delivered" per unit time is simply proportional to the volume of blood delivered to the tissues per unit time. We also assume for simplicity that among species being compared, tissue density is the same at all body sizes, such that mass and volume scale isometrically.

The most efficient transportation network (14, 15) for supplying metabolites is one that is *directed*—the flow of metabolites is always away from the source, and there is no large-scale backtracking. We first consider how the simple geometric properties of such a network allow us to express its capacity to deliver metabolites at minimal cost at different body sizes.

Consider a system in D dimensions ($D = 3$ in most cases of interest) having a volume, V, and mass, M, both proportional to L_p^D (L_p denotes the physical length of the system):

$$M \propto V \propto L_p^D \qquad [1]$$

The total metabolic rate of the organism, B, scales as

$$B \propto L_p^D E, \qquad [2]$$

where E is the rate of energy use per unit volume (and scales in the same manner as does mass-specific metabolic rate or metabolic intensity). Our aim is to determine how E scales with M.
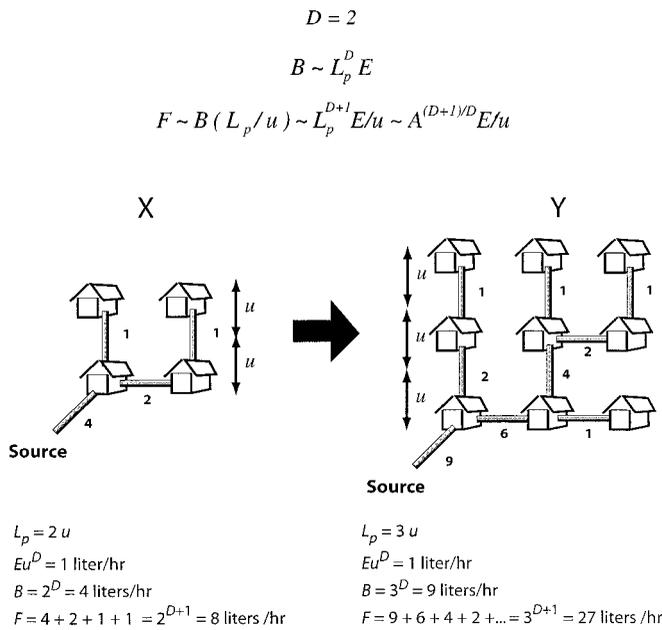
Consider a transportation network delivering metabolites to a set of nodes. Let u be the characteristic length scale, independent of body mass, separating neighboring nodes of the network. The local flow rate in the network is the amount of mass transported past a given location per unit time. An efficient transportation system (14) delivering metabolites satisfies

$$F \propto (L_p/u)B, \qquad [3]$$

where F is the sum of all individual flow rates. Physically, for any directed transportation network L_p/u is the mean distance from the source to the nodes, measured in terms of the basic length unit, u. Eq. **3** has been proven as a mathematical theorem (14). Fig. 1 illustrates the theorem and its consequences for directed transportation networks.

$$D = 2$$

$$B \sim L_p^D E$$

$$F \sim B(L_p/u) \sim L_p^{D+1} E/u \sim A^{(D+1)/D} E/u$$

X   Y

Source

Source

$L_p = 2\,u$

$Eu^D = 1$ liter/hr

$B = 2^D = 4$ liters/hr

$F = 4 + 2 + 1 + 1 = 2^{D+1} = 8$ liters /hr

$L_p = 3\,u$

$Eu^D = 1$ liter/hr

$B = 3^D = 9$ liters/hr

$F = 9 + 6 + 4 + 2 + \ldots = 3^{D+1} = 27$ liters /hr

Fig. 1. Example of general scaling properties of directed transportation networks. The figure illustrates in a schematic way the mathematical theorem proven in ref. 14, a nonbiological case in which the volume of the transport network is *not* constrained to scale isometrically with the volume that it is supplying (see text). This theorem can be directly applied to the study of network allometry (14, 19). Symbols are as defined in the text. Our example network consists of water ("blood") being transported from a source ("the heart") at a constant rate to each house of a space-filling neighborhood of identical houses ("all parts of the body"). For ease of illustration, we have represented everything as if it occurred in two dimensions ($D = 2$), but the argument and theorem apply in the same way to the three-dimensional case. Consider two neighborhoods of different size (X and Y, where the number of houses is proportional to their respective areas, A). The amount of water each pipe must carry per unit time is indicated adjacent to the respective pipe in the figure. The total flow rate of water (F) in a network must scale as $A^{(D+1)/D}$ to supply each house at a constant rate independent of the size of the neighborhood. Thus the quantity of the circulating water increases superlinearly with neighborhood size. A more detailed explanation is as follows: Neighborhoods X and Y are of different sizes (characterized by different physical lengths, L_p). To make units explicit, we standardize our length scale to be in units of u, a length based on the size of a typical house and its immediate environs. The area (or volume) associated with a house is u^D. The amount of water leaving the source per unit time and thence arriving at the houses constitutes the total "metabolic rate" (B) of each neighborhood. All of the houses are identical, and in each neighborhood they consume the same amount of water per unit time per house (Eu^D, where E is the rate of water use per unit area). (Note that in the biological case, E would depend on the size of the "neighborhood.") The sum of the individual flow rates (the amount of water circulating in all the pipes of a neighborhood per unit time), F, is proportional to B but with a proportionality constant equal to the mean distance from the source to all the houses, measured along the transportation route. This additional factor arises because the transportation system contains water bound to all houses, and each house places an additional requirement on F of an amount proportional to its distance from the source. For example, to satisfy the requirement that a house three units ($3\,u$) from the source is supplied by an amount Eu^D of water (over every unit time interval), it is necessary to ensure that there is a total amount $3\,Eu^D$ of water in the pipes on the way to the house that will not be used by any intervening house. The most efficient transportation system, characterized by the smallest F, is one in which this mean distance from the source to the houses is as small as possible. This happens when the transportation is directed away from the source toward the houses. In a D dimensional space, B scales as $L_p^D E$, where L_p is the physical length of the neighborhood, and the number of houses scales as $(L_p/u)^D$. Our theorem (14, 15) asserts that because the additional factor corresponding to the mean distance from the houses to the source must itself scale at least as L_p/u, F must scale at least as BL_p/u. This means that if $A = L_p^D$ is the conventional "size" of a neighborhood (its area if $D = 2$ or its volume if $D = 3$), then

For organisms, there is an important additional constraint that does not apply to many physical systems (such as the one described in Fig. 1). An organism's transport network must be contained entirely within the body, and it thus constitutes part of the body volume and mass. It is observed empirically that, among mammals, blood mass is a constant proportion (6–7%) of body mass across more than 5 orders of magnitude in size (4, 5). From a functional, and developmental point of view, this is understandable. If in organisms the circulatory system were to scale as does the network depicted in Fig. 1 (to maintain constant nutrient delivery rate per unit volume of body mass, i.e., a constant E), the blood volume, assumed in ref. 14 to be proportional to F, would scale with body mass with an exponent greater than 1 (4/3 when $D = 3$). Large species would then have a great deal of blood per unit volume of tissue, whereas small species would have relatively very little. In fact, if we start with an ordinary 50-g mouse (assuming blood mass to be 7% of body mass) and scale its blood volume to larger body sizes by using the exponent of 4/3, we see that mammals of as little as 160 kg in mass would have to be composed entirely of blood. Clearly, biological function would be compromised long before this size was reached. Even relatively small changes of body size would entail radical changes in composition of the body's tissues, not what is generally observed. Indeed, in all groups of organisms, tissue composition is conserved as much as possible across body sizes. We assume then that generally the volume and mass of the blood will scale approximately isometrically with body mass and be proportional to L_p^3, as is the case in mammals. The main consideration in the present context is that mass-specific meta-bolic demand must change to match the way the capacity of the network scales when it is subject to this constraint.

Because under the foregoing constraint the total mass of the blood scales isometrically with the total mass of the body, we can replace the blood mass with body mass in any expression without changing the body-mass scaling behavior of any dependent quantity. Accordingly, we now use the body mass, M, and F (the total blood flow, which has units of mass/time) to define a quantity that has the dimensions of inverse time and is purely a property of the geometry of the transportation network,

$$r_1 = F/M. \qquad [4]$$

The quantity r_1 is a measure of the total flow rate of the metabolites per unit mass of the organism. Combining Eqs. **1–4**, we find

$$r_1 \propto E L_p / u. \qquad [5]$$

The usefulness of this quantity will become apparent shortly. Note that r_1 is *not* the measured rate of *delivery* of metabolites. Rather, r_1 represents (in its scaling behavior with respect to body mass M) the way that the geometry of an efficient network in an organism changes with body mass (and in so doing imposes a

$$F \propto A^{(D+1)/D} E/u$$

and the total flow rate per unit area in the system thus increases with neighborhood size; this can be readily confirmed from the figure. (An even simpler example is the case of a one-dimensional network of length L_p, F is proportional to $1+2+3+4+\cdots+L_p$. This sum scales as L_p^2 in accord with the above equation.) It can be proven exactly that this result is independent of whether or not one has an underlying network, as long as the flow is "radially" outward from the source to the sinks (17, 18). Furthermore, the result is the same whether or not the houses are terminal units and whether or not they lie on a regular network (as shown in Fig. 1). Also, the scaling behavior holds in the large length scale limit independently of whether the source is at the periphery of the neighborhood, as shown in the figure, or whether it is somewhere within the neighborhood.

cost, in terms of additional mass of blood per unit increase in body mass, required to maintain constant delivery rates). For a given E, a larger r_1 means that this specific cost for the network to deliver metabolites at a constant rate to the tissues is greater. (Indeed, an inefficient undirected network would lead to even larger specific costs; ref. 14). The value of r_1 is proportional to the local demand for metabolites (E), whereas the scaling properties of an efficient transportation network (see Fig. 1) lead to the additional factor of L_p/u in Eq. 5, which obviously itself scales as a power of body mass. If organisms of different sizes exhibit the same r_1, they have the same specific delivery costs, and such a case implies that E must vary across body size in such a way as to compensate for increasing L_p/u. Note that r_1 is inversely proportional to u and therefore depends on the details of network geometry.

At this point, we have an expression for a rate, associated with the scaling of the capacity of the network to supply metabolites, that depends on E, the metabolic demand per unit volume. If we now can express the geometry of the scaling of metabolic demand in a comparable form, we can determine the scaling of metabolism in the case where demand and supply are matched.

Accordingly, we now switch our perspective from the supply network to the demand locations, i.e., to what happens at the nodes of the transportation network, where the metabolites are being delivered. Our goal is to work out the scaling behavior of the rate, r_2, with which the metabolites are consumed or taken in at the level of the tissues. To accomplish this, let us define a service volume of spatial extent l_s and volume l_s^D, whose total consumption is 1 metabolite (or energy) unit per unit time, independent of body mass. Thus

$$El_s^D \propto 1. \qquad [6]$$

Noting that 1 time unit is associated with the length l_s, the time scale corresponding to the length scale u must be proportional to u/l_s, so that the "demand" rate (measured in inverse time units) is given by

$$r_2 \propto l_s/u. \qquad [7]$$

Although l_s would be expected to vary with body mass, the rate of consumption of metabolites in a service volume is invariant with body mass. On dimensional grounds, the quantity r_2 is the unique rate at the demand locations that is inversely proportional to u.

Because r_2, like r_1, is inversely proportional to u, the nonuniversal dependence on u can be eliminated on taking the ratio of r_1 and r_2:

$$\frac{r_1}{r_2} = \frac{EL_p}{l_s}. \qquad [8]$$

Combining Eqs. 1 and 2, we obtain $B \propto ME$ and therefore $B^{(D+1)/D} \propto M^{(D+1)/D}E^{(D+1)/D} = MEM^{1/D}E^{1/D} \propto M(EL_p/l_s)$ on using Eqs. 1 and 6, which in turn equals Mr_1/r_2 (see Eq. 8), so that

$$B \propto \left(M\frac{r_1}{r_2} \right)^{D/(D+1)}, \qquad [9]$$

which is independent of the length scale u.

We have considered two distinct processes associated with metabolism: the delivery of metabolites through a network with an associated rate r_1 and the demand for the delivered metabolites at a rate r_2. Maintaining a match between these two rates across body sizes would require that both rates scale with body mass in the same manner, i.e., if $r_1 \propto M^{s_1}$ and $r_2 \propto M^{s_2}$, that $s_1 = s_2$. If this were not true, under changes of body mass either the

supply of the metabolites would exceed the demand or vice versa. Although a small degree of mismatch of the exponents associated with the two rates may be tolerable over a limited range of body mass, the harmonious matching of r_1 and r_2 would be required to maintain efficiency over a large range of body mass. Such matching of the functional components of physiological systems is expected and commonly observed (16). When $s_1 = s_2$, one finds straightforwardly the 3/4 scaling of metabolism and associated relationships: $B \propto M^{D/(D+1)}$, $r_1 \propto r_2 \propto l_s \propto M^{1/D(D+1)}$ and $E \propto M^{-[1/(D+1)]}$. Physiological time τ_0 may be defined in the usual manner as being inversely proportional to E, leading to the result that $\tau_0 \propto M^{1/(D+1)}$. From the above, a large number of observed scaling relationships widely discussed in the literature may be straightforwardly derived (5).

The dominant effect of the scaling properties of the circulatory system on the scaling of metabolism may seem surprising, given that the rate of supply of metabolites to the tissues is only one of a series of physiological and biochemical steps involved in metabolism (16). However, here we are concerned only with how metabolic rates may scale interspecifically and not with how they are controlled within an individual. Intracellular processes and properties—including the rates of chemical reactions in organelles, the function and concentration of enzymes, and the strength of chemical bonds—are most unlikely to exhibit necessary changes in direct response to the overall size of the organism. In contrast, the capacity of the circulatory system, as we have shown here, necessarily scales with body mass. The correspondence between the scaling exponent for the capacity of the circulatory system and that observed for overall metabolism allows us to draw two conclusions: the rates of intracellular metabolic-related processes conform roughly to the scaling of the supply network and exert little, if any, net effect on the scaling of overall metabolism, and the overall metabolic demand per unit body mass is indeed ordinarily matched to the capacity of the supply network. The latter implies that $s_1 = s_2$ in most groups of similar organisms.

We have not specified the physiological mechanisms that result in the maintenance of approximate equality of the scaling exponents of r_1 and r_2, and the actual mechanisms may differ among different types of organisms. The scaling behavior of r_2 refers to the scaling of the total metabolic demand per unit body mass, and there are multiple ways to alter this value, including changes in the rates of metabolic chemical reactions and changes in anatomy at the cell and tissue levels.

Some degree of inequality between the s values may nevertheless be maintained—on either short or long time scales—if other physiological processes can compensate for the resulting inefficiency. We can make some general statements about variations in the way that the supply and demand rates would be expected to scale and the consequences for metabolic scaling exponents. Both s_1 and s_2 ought to be nonnegative, because neither the supply nor the demand rate can decrease as the mass increases. On general grounds (16), it is wasteful for a biological organism to be characterized by a supply rate greater than the demand rate, although this would still permit function. These considerations imply that ordinarily $s_1 \leq s_2$. The two most interesting cases, therefore, correspond to: $s_1 = s_2 = [1/D(D+1)]$, where the 3/4-power scaling of total metabolic rate is obtained and there is a perfect balance between supply and demand as derived above; and the situation when $s_1 = 0$ corresponds to an r_1 independent of body mass. (This situation could arise, for example, if additional tissue that needed to be supplied with metabolites were added to the body, but with no change in the geometry of the network—for example, during ontogenetic growth or among a group of closely related species.) In this latter case, Eqs. 1, 2, 5, 6, 7, and 9 may readily be used to show that the 2/3-power scaling of total metabolic rate emerges without any reference to the surface to volume ratio. Specifically,

$B \propto M^{(D-1)/D}$, $r_2 \propto l_s \propto M^{1/(D^2)}$ and $E \propto M^{-(1/D)}$. Physiological time in this case turns out to scale as $\tau_0 \propto M^{1/D}$. Thus, the exponent associated with the scaling of the total metabolic rate with mass is predicted to lie usually between 2/3 and 3/4, with the latter value being preferred because of the supply–demand balance.

Our theory ought to be generally applicable and should describe the major features of the biological scaling of a large variety of organisms, because it is based on a minimal set of essential biological details. Because the theorem pertaining to the scaling of flow in directed networks (Eq. **3**) is valid for directed flow in general (14, 15, 17, 18; Fig. 1), and its main implication is robust to geometrical fluctuations (19), our results should also hold for types of organisms that may not possess obvious branching transportation networks, such as unicellular species.

The two foundations of our theory are the scaling properties of optimal transportation networks and the balance between supply and demand. The latter, as introduced and developed in this paper, permits us to describe the specific but very simple conditions under which quarter-power scaling of organismic metabolism would be expected. When the scaling of demand balances the scaling of supply, the commonly observed allometric scaling laws follow directly from the geometric properties of the transportation network. Furthermore, our theory provides a framework for understanding deviations from the 3/4 exponent: supply–demand scaling imbalances stemming from inefficiency or permitted by compensating physiological mechanisms.

1. McMahon, T. A. & Bonner, J. T. (1983) *On Size and Life* (Freeman, New York).
2. Bonner, J. T. (1983) *The Evolution of Complexity by Means of Natural Selection* (Princeton Univ. Press, Princeton).
3. Peters, R. H. (1983) *The Ecological Implications of Body Size* (Cambridge Univ. Press, Cambridge, U.K.).
4. Schmidt-Nielsen, K. (1984) *Scaling: Why Is Animal Size So Important?* (Cambridge Univ. Press, Cambridge, U.K.).
5. Calder, W. A., III (1984) *Size, Function, and Life History* (Harvard Univ. Press, Cambridge, MA).
6. Enquist, B. J., Brown, J. H. & West, G. B. (1998) *Nature (London)* **395,** 163–165.
7. McNab, B. K. (1988) *Q. Rev. Biol.* **63,** 25–54.
8. Dodds, P. S., Rothman, D. H. & Weitz, J. S. (2001) *J. Theor. Biol.* **209,** 9–27.
9. Feldman, H. A. & McMahon, T. A. (1983) *Respir. Physiol.* **52,** 149–163.
10. Damuth, J. (1998) *Nature (London)* **395,** 115–116.
11. West, G. B., Brown, J. H. & Enquist, B. J. (1997) *Science* **276,** 122–126.
12. West, G. B., Brown, J. H. & Enquist, B. J. (1999) *Science* **284,** 1677–1679.
13. West, G. B., Brown, J. H. & Enquist, B. J. (1999) *Nature (London)* **400,** 664–667.
14. Banavar, J. R., Maritan, A. & Rinaldo, A. (1999) *Nature (London)* **399,** 130–131.
15. Banavar, J. R., Maritan, A. & Rinaldo, A. (2000) *Nature (London)* **408,** 160.
16. Weibel, E. R., Taylor, C. R. & Bolis, L., eds. (1998) *Principles of Animal Design: The Optimization and Symmorphosis Debate* (Cambridge Univ. Press, Cambridge, MA).
17. Dreyer, O. (2001) *Phys. Rev. Lett.* **87,** 038101.
18. Dreyer, O. & Puzio, R. (2001) *J. Math. Biol.* **43,** 144–156.
19. Maritan, A., Rigon, R., Banavar, J. R. & Rinaldo, A. (2002). *Geophys. Res. Lett.* **29,** 31–34.

APPLIED MATHEMATICS

ECOLOGY