A Common Scaling Rule for Abundance, Energetics, and Production of Parasitic and Free-Living Species

Ryan F. Hechinger, *et al.*

Science **333**, 445 (2011);
DOI: 10.1126/science.1204337

long necks and tails could have facilitated heat dissipation by increasing their surface area (*37*). Overall, our data are most consistent with the hypothesis that sauropods sustained high metabolic rates during ontogeny to reach their gigantic size so rapidly, but that in maturity a combination of physiological and behavioral adaptations and/or a slowing of metabolic rate prevented problems with overheating and avoided excessively high body temperatures (*18*, *36*). An unresolved question is whether such adaptations could have compensated for the high internal heat production associated with endothermy, or whether large adult sauropods must have had both heat-dissipating adaptations and a low basal metabolism to maintain body temperatures in the 36° to 38°C range that we have measured.

References and Notes

1. L. S. Russell, *J. Paleontol.* **39**, 497 (1965).
2. A. J. de Ricqlès, *Evol. Theory* **1**, 51 (1974).
3. R. T. Bakker, *Nature* **238**, 81 (1972).
4. R. T. Bakker, in *A Cold Look at the Warm-Blooded Dinosaurs,* R. D. K. Thomas, E. C. Olson, Eds. (Westview, Boulder, 1980), pp. 351–462.
5. J. O. Farlow, in *The Dinosauria,* D. B. Weishampel, P. Dodson, H. Osmólska, Eds. (Univ. California Press, Berkeley, 1990), pp. 43–55.
6. M. P. O'Conner, P. Dodson, *Paleobiology* **25**, 341 (1999).
7. F. Seebacher, G. C. Grigg, L. A. Beard, *J. Exp. Biol.* **202**, 77 (1999).
8. F. Seebacher, *Paleobiology* **29**, 105 (2003).
9. J. F. Gillooly, A. P. Allen, E. L. Charnov, *PLoS Biol.* **4**, e248 (2006).
10. H. Pontzer, V. Allen, J. R. Hutchinson, *PLoS ONE* **4**, e7783 (2009).
11. G. M. Erickson, *Trends Ecol. Evol.* **20**, 677 (2005).
12. J. Ruben, A. Leitsch, W. Hillenius, N. Geist, T. Jones, in *The Complete Dinosaur,* J. O. Farlow, M. K. Brett-Surman, Eds. (Indiana Univ. Press, Bloomington, 1997), pp. 505–518.
13. R. E. Barrick, W. J. Showers, *Science* **265**, 222 (1994).
14. H. C. Fricke, R. R. Rogers, *Geology* **28**, 799 (2000).
15. R. Amiot *et al.*, *Earth Planet. Sci. Lett.* **246**, 41 (2006).
16. K. Padian, J. R. Horner, in *The Dinosauria,* D. B. Weishampel, P. Dodson, H. Osmólska, Eds. (Univ. California Press, Berkeley, 2004), pp. 660–671.
17. A. Chinsamy, W. J. Hillenius, in *The Dinosauria,* D. B. Weishampel, P. Dodson, H. Osmólska, Eds. (Univ. California Press., Berkeley, 2004), pp. 643–659.
18. P. M. Sander *et al.*, *Biol. Rev. Camb. Philos. Soc.* **86**, 117 (2011).
19. T. J. Case, *Q. Rev. Biol.* **53**, 243 (1978).
20. A. J. de Ricqlès, in *A Cold Look at the Warm-Blooded Dinosaurs,* R. D. K. Thomas, E. C. Olson, Eds. (Westview, Boulder, 1980), pp. 103–139.
21. P. M. Sander, *Paleobiology* **26**, 466 (2000).
22. N. Klein, P. M. Sander, *Paleobiology* **34**, 247 (2008).
23. A. E. Dunham, K. L. Overall, W. P. Porter, C. A. Forster, in *Paleobiology of the Dinosaurs, GSA Special Paper 238,* J. O. Farlow, Ed. (Geological Society of America, Boulder, 1989), pp. 1–21.
24. J. R. Spotila, M. P. O'Conner, P. Dodson, F. V. Paladino, *Mod. Geol.* **16**, 203 (1991).
25. R. M. Alexander, *Palaeontology* **41**, 1231 (1998).
26. F. V. Paladino, M. P. O'Conner, J. R. Spotila, *Nature* **344**, 858 (1989).
27. Materials and methods are available as supporting material on *Science* Online.
28. P. Ghosh *et al.*, *Geochim. Cosmochim. Acta* **70**, 1439 (2006).
29. R. A. Eagle *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **107**, 10377 (2010).
30. F. Seebacher, R. M. Elsey, P. L. Trosclair 3rd, *Physiol. Biochem. Zool.* **76**, 348 (2003).
31. K. J. Dennis, D. P. Schrag, *Geochim. Cosmochim. Acta* **74**, 4110 (2010).
32. M. J. Kohn, T. E. Cerling, in *Reviews in Mineralogy and Geochemistry. Phosphates: Geochemical, Geobiological, and Materials Importance*, vol. 48, M. J. Kohn, J. Rakovan, J. M. Hughes, Eds. (Mineralogical Society of America and Geochemical Society, Washington, DC, 2002), pp. 455–488.
33. H. C. Fricke, R. R. Rogers, R. Backlund, C. N. Dwyer, S. Echt, *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **266**, 13 (2008).
34. A. Clarke, P. Rothery, *Funct. Ecol.* **22**, 58 (2008).
35. M. J. Wedel, *Paleobiology* **29**, 243 (2003).
36. P. M. Sander, M. Clauss, *Science* **322**, 200 (2008).
37. E. H. Colbert, *Am. J. Sci.* **293**, (A), 1 (1993).

Acknowledgments: R.A.E. and J.M.E. are funded by National Science Foundation grant EAR-1024929. T.T. is funded by Deutsche Forschungsgemeinschaft (DFG) grants TU 148/2-1 and TU 148/4-1. A.K.T. is supported by the Division of Physical Sciences at UCLA and National Science Foundation grant EAR-0949191. We thank the Utah Museum of Natural History, the Oklahoma Museum of Natural History, the Berlin Museum für Naturkunde, the Aathal Dinosaurier Museum, D. Chure at Dinosaur National Monument, and the Wyoming Dinosaur International Society for the provision of specimens for analysis. We thank D. Schwarz-Wings for specimen identification and D. Goodreau for curatorial assistance. We also thank the stable isotope laboratories at Tübingen and Lyon for conducting phosphate $\delta^{18}O$ measurements. Lastly, thanks to J. Gillooly for provision of published data. This is contribution number 113 of the DFG Research Unit 533 "Biology of the Sauropod Dinosaurs: The Evolution of Gigantism." R.A.E. designed the project and experiments, performed the experiments, analyzed the data, and wrote the manuscript. J.M.E. and A.K.T. assisted with experimental design and data interpretation. Clumped isotope analyses were conducted in the lab of J.M.E. T.T. provided samples for the study, conducted phosphate oxygen isotope analysis, and assisted with data interpretation. T.S.M. assisted with isotope measurements and petrographic analyses. H.C.F., M.C., and R.L.C. provided samples for analysis and insights into the provenance of samples and data interpretation.

Supporting Online Material

www.sciencemag.org/cgi/content/full/science.1206196/DC1
Materials and Methods
SOM Text
Figs. S1 to S6
Tables S1 to S10
References (*38*–*103*)

28 March 2011; accepted 3 June 2011
Published online 23 June 2011;
10.1126/science.1206196

A Common Scaling Rule for Abundance, Energetics, and Production of Parasitic and Free-Living Species

Ryan F. Hechinger,[1]* Kevin D. Lafferty,[1,2] Andy P. Dobson,[3,4] James H. Brown,[5] Armand M. Kuris[1]

The metabolic theory of ecology uses the scaling of metabolism with body size and temperature to explain the causes and consequences of species abundance. However, the theory and its empirical tests have never simultaneously examined parasites alongside free-living species. This is unfortunate because parasites represent at least half of species diversity. We show that metabolic scaling theory could not account for the abundance of parasitic or free-living species in three estuarine food webs until accounting for trophic dynamics. Analyses then revealed that the abundance of all species uniformly scaled with body mass to the −¾ power. This result indicates "production equivalence," where biomass production within trophic levels is invariant of body size across all species and functional groups: invertebrate or vertebrate, ectothermic or endothermic, and free-living or parasitic.

General ecological theory should apply to all species, and thus should include the parasites that represent at least half of species diversity (*1*–*3*). A goal of the metabolic theory of ecology is to broadly explain and predict local species abundance by considering how metabolic rate scales with body size and temperature (*4*, *5*). Although studies have documented the scaling of parasite abundance with body size within individual hosts (*6*, *7*), none have examined the scaling of parasites alongside co-occurring free-living species. This omission is potentially critical because, in addition to their great diversity, there are other factors indicating that the inclusion of parasites can test and refine general rules for abundance and body-size scaling.

Parasites differ from free-living consumers in ways that can violate assumptions made by current models of abundance and diversity. For instance, because parasites are smaller than their hosts, they invert consumer-resource body-size ratios, which are often assumed to be constant and larger than 1 (*4*, *8*–*10*). Further, parasites might be rarer than other small consumers, as they tend to occupy higher trophic levels to which the flow of resources is constrained by trophic

[1]Marine Science Institute and Department of Ecology, Evolution, and Marine Biology, University of California, Santa Barbara, CA 93106, USA. [2]Western Ecological Research Center, U.S. Geological Survey, Marine Science Institute, University of California, Santa Barbara, CA 93106, USA. [3]Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ 08544, USA. [4]Santa Fe Institute, Santa Fe, NM 87501, USA. [5]Department of Biology, University of New Mexico, Albuquerque, NM 87131, USA.

*To whom correspondence should be addressed. E-mail: hechinger@lifesci.ucsb.edu

transfer efficiency (the fraction of energy in resource populations that is converted into consumer populations) (*11*–*13*). On the other hand, parasites appear to grow and reproduce at higher rates than their free-living relatives (*14*), potentially reflecting higher assimilation and production efficiencies, which are major components of trophic transfer efficiency (*11*–*13*). These differences highlight the importance of simultaneously considering parasites and free-living species to develop empirical generalizations and theory concerning species abundance, body-size scaling, and trophic dynamics.

Among species, abundance typically decreases with increasing body size because larger individuals require more resources. Resource requirements parallel whole-organism metabolic rates, which increase with body size as M^{α}, where M is body mass and α is a scaling exponent with a positive sign. Consequently, when differently sized species have, on average, equal access to resources, population abundance N is predicted to scale as

$$N = iM^{-\alpha} \quad (1)$$

where i is a normalization constant (*4*, *5*, *15*). Body temperature also influences abundance, because metabolic rates can increase over a broad temperature range (*5*, *16*). All else being equal, abundance decreases with increasing temperature because each individual requires more resources (*4*). We can add temperature dependence, so that

$$N = iM^{-\alpha}D^{-1} \quad (2)$$

where D represents a dimensionless temperature-dependence term (*17*, *18*). This can be captured by an Arrhenius equation formulation that expresses exponential temperature effects relative to a standard: $\exp[(E/k)(T - T_o)/TT_o]$, where E is the activation energy, k is the Boltzmann constant, T is the body temperature, and T_o is the standard temperature (*11*, *16*). We can rearrange Eq. 2 to provide an estimate of "temperature-corrected" abundance (*4*, *17*): $N_{temp} = ND = iM^{-\alpha}$. Abundance–body size relationships are usually analyzed by linear regression after logarithmic transformation, so that $\log N_{temp} = \log i - \alpha \log M$, and the slope gives the exponent α. Because whole-organism metabolic rates across a broad range of multicellular organisms scale, on average, as $M^{3/4}$ (*5*, *16*, *19*, *20*), their abundance is predicted to scale, on average, as $M^{-3/4}$ (*4*, *8*–*10*, *21*, *22*).

Few studies (*23*, *24*) have quantified the scaling of local abundance for diverse groups of species that coexist in an ecosystem and span a wide range of body sizes and basic physiologies, and none have included parasites. Recent investigations of three estuarine food webs in California and Baja California provide data that permit such an analysis (*25*, *26*) (table S1). In all three estuaries, species abundance decreased by 11 orders of magnitude as body size increased by 11 orders of magnitude (Fig. 1, A to C). A single regression did not adequately describe the relationship, primarily because slopes for parasites (–0.50 to –0.63) were consistently shallower than the slopes characterizing free-living species (–1.26 to –1.36) (Fig. 1, A to C). Furthermore, parasites were consistently less abundant than free-living species of similar body size. Despite being two to three orders of magnitude smaller than the average free-living invertebrate species, the average parasite species was at least one order of magnitude less abundant. As expected, using N_{temp} to factor in the higher body temperature of birds relative to the other animals [characterized by ambient environmental temperatures (*11*)] provided very similar relationships (Fig. 1, D to F). Thus, in these analyses, the scaling of abundance with body size does not support a common scaling exponent, –¾ or otherwise, for parasitic and free-living species.

However, the above analysis does not account for the flow of energy among trophic levels. Inefficiencies in exploitation, assimilation, and production ensure that trophic transfer efficiency is less than 100% (*12*, *13*). Thus, fewer resources are available to higher trophic levels. Previous studies have added trophic transfer efficiency to scaling relationships by modifying the scaling exponent, assuming a particular transfer efficiency and that consumer-resource body-size ratios are fixed and larger than 1 (*4*, *8*–*10*, *21*, *27*, *28*). Free-living assemblages will sometimes violate this assumption, potentially explaining why adding trophic transfer efficiency in this manner performed no better than ignoring it did in a previous analysis of 121 food webs (*29*). Moreover, simultaneous consideration of parasitic and free-living species will always strongly violate assumptions of consistent consumer-resource body-size ratios or of a positive relationship between trophic level and body size.

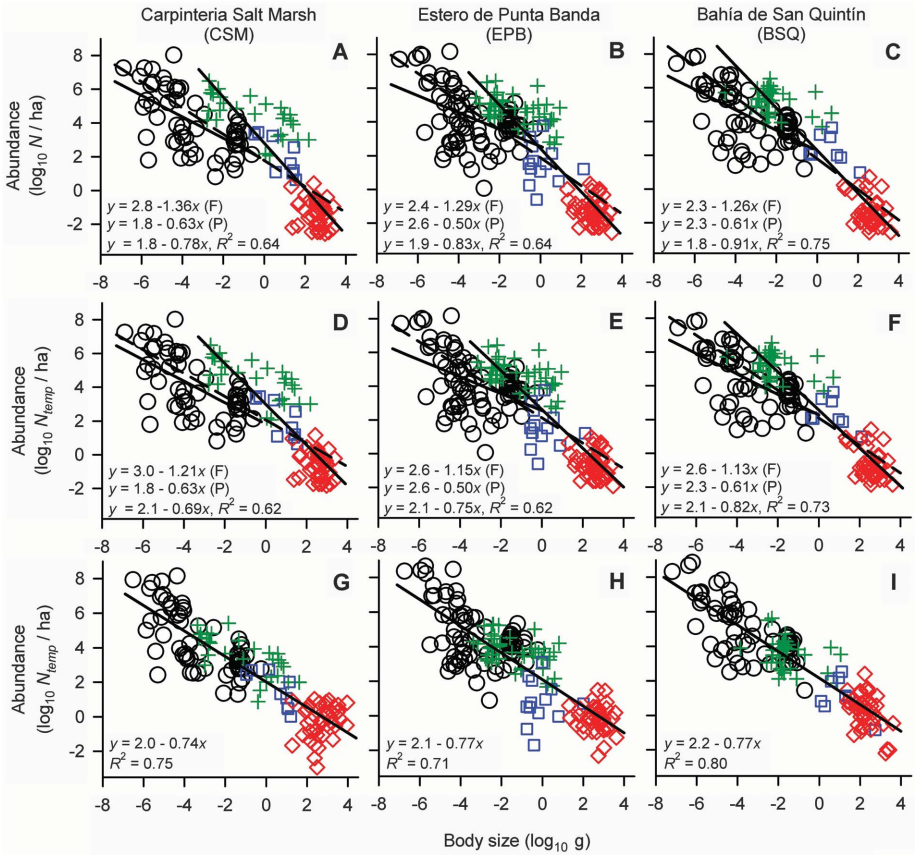

Fig. 1. Abundance as a function of body size for parasitic and free-living species in three estuaries: Carpinteria Salt Marsh (CSM), Estero de Punta Banda (EPB), and Bahía de San Quintin (BSQ). (**A** to **C**) Abundance versus body size reveals that a single regression line cannot adequately fit the data (general linear models: all interaction Ps < 0.0001; tables S2 and S3). Solid lines and top two equations give the slopes and intercepts for parasitic (P) and free-living (F) species; slope 95% confidence limits: CSM, ±0.14; EPB, ±0.13; BSQ, ±0.11. The broken lines, bottom equations, and R^2s pertain to pooled data. (**D** to **F**) Temperature-corrected abundance versus body size gives relationships very similar to those seen in (A) to (C), although bird abundance is shifted up by about half an order of magnitude, leading to slightly shallower slopes for free-living and pooled data. Lines and equations as in (A) to (C); slope 95% confidence limits: CSM, ±0.13; EPB, ±0.12; BSQ, ±0.10. (**G** to **I**) Temperature-corrected abundance versus body size, statistically controlling for trophic level (Fig. 3 and tables S4 and S5). The scaling slopes are all consistent with the –¾ predicted by metabolic scaling, as slightly modified for the distribution of the number of species along the body-size axis (*11*); slope 95% confidence limits: CSM, ±0.073; EPB, ±0.073; BSQ, ±0.063. The R^2 values represent partial R^2s for body size. Symbol key for all figures: circles, parasites; crosses, invertebrates; squares, fish; diamonds, birds.

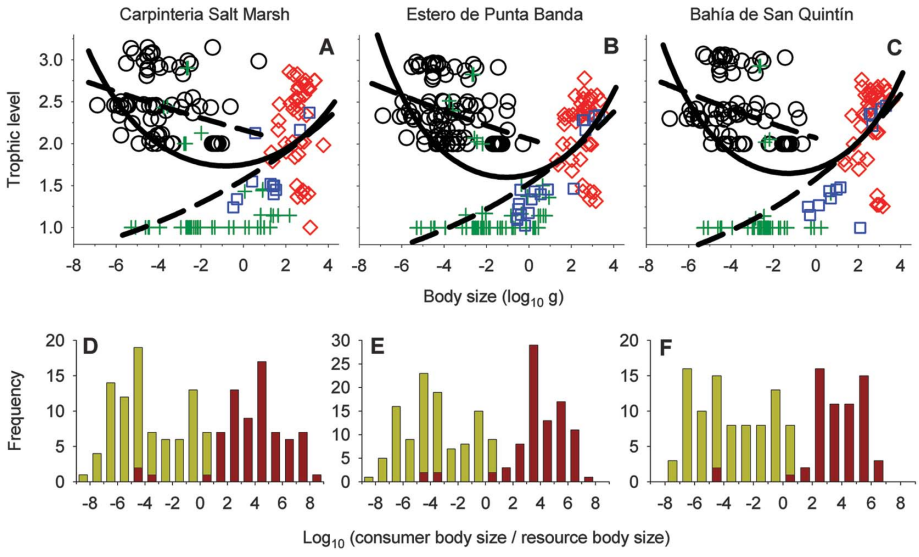

Fig. 2. Variation in trophic level with body size, and in consumer-resource body-size ratios, for parasitic and free-living species in three estuarine food webs. (**A** to **C**) Relationship between trophic level and body size. Dashed lines represent separate relationships for parasitic and free-living species (Poisson regressions, all interaction *P*s < 0.0001; tables S6 and S7), and solid lines represent significant curvilinear relationships for the two groups pooled (Poisson regressions, all quadratic term *P*s < 0.0001; tables S8 and S9). Symbols as in Fig. 1. (**D** to **F**) Frequency distributions of logged consumer-resource body-size ratios. Shaded portions of the histograms represent parasites and unshaded portions represent free-living consumers. Values less than 0 are for consumers that are smaller than their resources. These data show wide variation in consumer-resource body-size ratios, in contrast to the more constrained values observed when ignoring parasites.

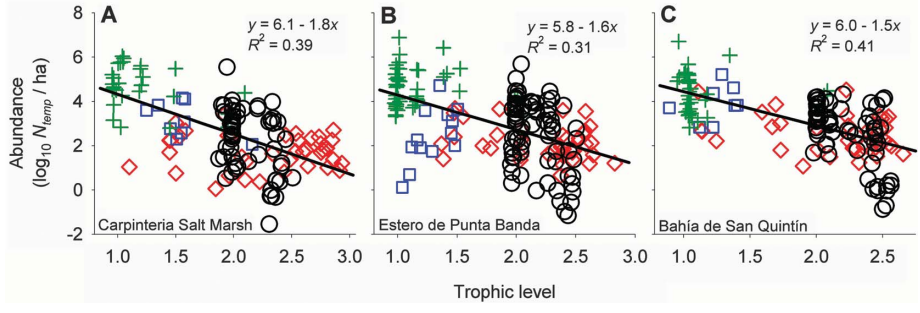

Fig. 3. (**A** to **C**) Abundance as a function of trophic level for parasitic and free-living species in three estuaries. Temperature-corrected abundance decreases with trophic level, as revealed by statistically controlling for body size (Fig. 1, G to I, and tables S4 and S5). The anti-log of the slope provides an estimate of ε, the overall trophic transfer efficiency in each ecosystem. Symbols as in Fig. 1.

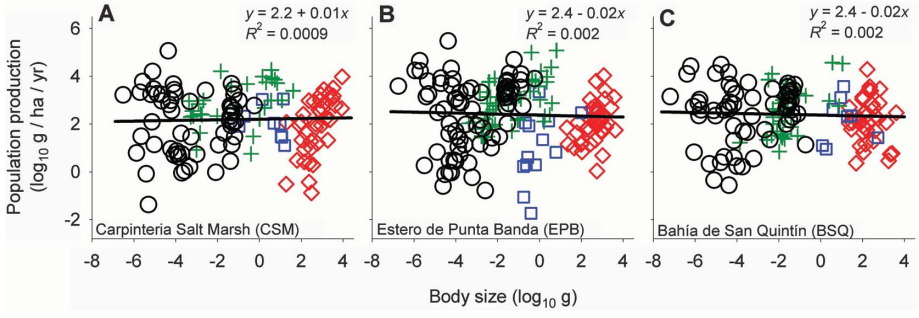

Fig. 4. (**A** to **C**) Population biomass production versus body size for parasitic and free-living species in three estuaries, statistically controlling for trophic level. The slopes of the fitted lines in each estuary are indistinguishable from zero (tables S10 and S11); 95% confidence limits: CSM, ±0.073; EPB, ±0.073; BSQ, ±0.063. Symbols as in Fig. 1.

Indeed, in the estuaries studied here, the relationship between trophic level and body size is U-shaped when parasites are included (Fig. 2, A to C) (*11*). Further, including parasites approximately doubled the range of observed consumer-resource body-size ratios because parasites, in contrast to typical free-living consumers, are much smaller than their resources (Fig. 2, D and E). This highlights the need to incorporate trophic transfer efficiency independent of body size, or of any assumed body-size associations, to derive broadly applicable and realistic scaling relationships. Although rarely done [e.g., (*4*, *27*, *28*)], we can use a separate, multiplicative term to capture the loss of energy among trophic levels. We can modify Eq. 2 to incorporate the exponential decrease of abundance with increasing trophic level (L) as

$$N_{\text{temp}} = iM^{-\alpha}\varepsilon^{L} \quad (3)$$

where ε is trophic transfer efficiency (a proportion) and basal trophic level = 0 (*11*). Linearizing Eq. 3 by log transformation gives

$$\log N_{\text{temp}} = \log i - \alpha \log M + \log \varepsilon \cdot L \quad (4)$$

Equation 4 can be analyzed directly with a general linear model that incorporates L and log M as predictor variables and provides empirical estimates of the scaling exponent and trophic transfer efficiency. After controlling for body size, as predicted, abundance decreased with increasing trophic level (Fig. 3, A to C). The estimates of average transfer efficiency across all species were $\varepsilon \approx 0.025$, toward the low end of the range typically reported (*12*, *13*, *30*). These estimates of ecosystem-wide transfer efficiency may be accurate, but they may also be a consequence of the theoretical assumption that only the bottom-up process of resource supply constrains abundance. The effect of trophic level may also include top-down effects of consumers (predators or parasites) on resource (prey or host) abundance, which can be explored by future research.

The use of Eq. 4 to incorporate trophic dynamics revealed a uniform ecosystem-wide scaling of abundance with body size in all three estuaries (Fig. 1, G to I). The relationships no longer differed between parasites and free-living species (table S4). Further, the slopes of the uniform abundance versus body-size relationships were all very close to –¾, as predicted by a ¾ scaling of metabolic rate with body size. Hence, after accounting for temperature, trophic level, and trophic transfer efficiency, a single line consistently explained abundance as a function of body size across diverse taxonomic and functional groups.

For physiologically similar multicellular organisms, a –¾ scaling of abundance with body size implies the average "energetic equivalence" of differently sized species because a single line describes the average $M^{3/4}$ scaling of whole-organism metabolic rates (*4*, *15*, *22*, *31*, *32*). In these cases, the population energy flux F (the product of individual metabolic rate and population abundance) scales invariant of body size, as $M^{3/4}M^{-3/4} = M^{0}$ (fig. S1). However, a single line

does not adequately describe the $M^{3/4}$ scaling of whole-organism metabolism for the species in our study because they span different physiological groups with different normalization constants (*4*, *16*) (fig. S1). Hence, the uniform abundance scaling documented here across all species indicates that, at any particular trophic level, populations of similarly sized species in different physiological groups flux different amounts of energy: endotherms > vertebrate ectotherms > parasitic or free-living invertebrates (fig. S1).

The uniform scaling of abundance found here has another general implication—that of "production equivalence." Specifically, species at the same trophic level produce biomass at the same average rate across all body sizes and functional groups. This occurs because, in contrast to metabolic rates, a single line can describe the $M^{3/4}$ scaling of individual biomass production, P_{ind}, for organisms of different physiological groups (*31*) (fig. S1). Consequently, the population production rate equals $P_{pop} = P_{ind}N$, which scales as $M^{3/4}M^{-3/4} = M^{0}$. Indeed, estimating population production for the species in the three estuaries supports the existence of this invariant biomass production with body size (Fig. 4 and fig. S1) (*11*). Thus, although population energy flux (and, consequently, demand on resources) may vary among physiological groups, opposing differences in production efficiency among these groups cause population biomass production to scale invariant of body size across all groups. Because production reflects biomass availability to consumers, production equivalence indicates a comparable ecological relevance for any single species within a trophic level, regardless of body size or functional group affiliation: invertebrate or vertebrate, ectotherm or endotherm, free-living or parasitic.

Accommodating parasitic and free-living species into a common framework highlights the utility of Eq. 3 to incorporate body size, temperature, and food-web information into ecological scaling theory in a simple and generally applicable way. Equations 3 and 4 may allow testing of the generality of the findings documented here for any ecosystem and any form of life.

References and Notes

1. P. W. Price, *Evolutionary Biology of Parasites* (Princeton Univ. Press, Princeton, NJ, 1980).
2. T. de Meeûs, F. Renaud, *Trends Parasitol.* **18**, 247 (2002).
3. A. P. Dobson, K. D. Lafferty, A. M. Kuris, R. F. Hechinger, W. Jetz, *Proc. Natl. Acad. Sci. U.S.A.* **105** (suppl. 1), 11482 (2008).
4. J. H. Brown, J. F. Gillooly, A. P. Allen, V. M. Savage, G. B. West, *Ecology* **85**, 1771 (2004).
5. R. H. Peters, *The Ecological Implications of Body Size* (Cambridge Univ. Press, Cambridge, 1983).
6. P. Arneberg, A. Skorping, A. F. Read, *Am. Nat.* **151**, 497 (1998).
7. S. Morand, R. Poulin, *Evol. Ecol. Res.* **4**, 951 (2002).
8. J. H. Brown, J. F. Gillooly, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 1467 (2003).
9. S. Jennings, S. Mackinson, *Ecol. Lett.* **6**, 971 (2003).
10. D. C. Reuman, C. Mulder, D. Raffaelli, J. E. Cohen, *Ecol. Lett.* **11**, 1216 (2008).
11. See supporting material on *Science* Online.
12. R. L. Lindeman, *Ecology* **23**, 399 (1942).
13. D. G. Kozlovsky, *Ecology* **49**, 48 (1968).
14. P. Calow, *Parasitology* **86**, 197 (1983).
15. J. Damuth, *Nature* **290**, 699 (1981).
16. J. F. Gillooly, J. H. Brown, G. B. West, V. M. Savage, E. L. Charnov, *Science* **293**, 2248 (2001).
17. A. P. Allen, J. H. Brown, J. F. Gillooly, *Science* **297**, 1545 (2002).
18. W. R. Robinson, R. H. Peters, J. Zimmermann, *Can. J. Zool.* **61**, 281 (1983).
19. M. Kleiber, *Hilgardia* **6**, 315 (1932).
20. A. M. Hemmingsen, *Repts. Steno. Hosp. Copenhagen* **9**, 7 (1960).
21. H. Cyr, in *Scaling in Biology*, J. H. Brown, G. B. West, Eds. (Oxford Univ. Press, Oxford, 2000), pp. 267–295.
22. J. Damuth, *Biol. J. Linn. Soc. London* **31**, 193 (1987).
23. H. Cyr, J. A. Downing, R. H. Peters, *Oikos* **79**, 333 (1997).
24. J. E. Cohen, T. Jonsson, S. R. Carpenter, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 1781 (2003).
25. R. F. Hechinger *et al.*, *Ecology* **92**, 791 (2011).
26. A. M. Kuris *et al.*, *Nature* **454**, 515 (2008).
27. T. D. Meehan, *Ecology* **87**, 1650 (2006).
28. B. J. McGill, *Am. Nat.* **172**, 88 (2008).
29. D. C. Reuman *et al.*, *Adv. Ecol. Res.* **41**, 1 (2009).
30. D. Baird, J. M. Mcglade, R. E. Ulanowicz, *Philos. Trans. R. Soc. B* **333**, 15 (1991).
31. S. K. M. Ernest *et al.*, *Ecol. Lett.* **6**, 990 (2003).
32. S. Nee, A. F. Read, J. J. D. Greenwood, P. H. Harvey, *Nature* **351**, 312 (1991).

Acknowledgments: We thank S. Sokolow, J. McLaughlin, J. Childress, and J. Damuth for discussion or comments on the manuscript. Supported by NSF/NIH EID grant DEB-0224565 and by CA Sea Grant R/OPCENV-01. The analyses in this manuscript used data published in Hechinger *et al.* (*25*), available at Ecological Archives (accession no. E092-066).

Supporting Online Material

www.sciencemag.org/cgi/content/full/333/6041/445/DC1
Materials and Methods
Figs. S1 and S2
Tables S1 to S11
References (*33*–*48*)

15 February 2011; accepted 27 May 2011
10.1126/science.1204337

Terraces in Phylogenetic Tree Space

Michael J. Sanderson,[1]* Michelle M. McMahon,[2] Mike Steel[3]

A key step in assembling the tree of life is the construction of species-rich phylogenies from multilocus—but often incomplete—sequence data sets. We describe previously unknown structure in the landscape of solutions to the tree reconstruction problem, comprising sometimes vast "terraces" of trees with identical quality, arranged on islands of phylogenetically similar trees. Phylogenetic ambiguity within a terrace can be characterized efficiently and then ameliorated by new algorithms for obtaining a terrace's maximum-agreement subtree or by identifying the smallest set of new targets for additional sequencing. Algorithms to find optimal trees or estimate Bayesian posterior tree distributions may need to navigate strategically in the neighborhood of large terraces in tree space.

Phylogenetic tree space, the collection of all possible trees for a set of taxa, grows exponentially with the number of taxa, creating computational challenges for phylogenetic inference (*1*). Nonetheless, phylogenetic trees and comparative analyses based on them are growing larger, with several exceeding 1000 species [e.g., (*2*)] and a recent one exceeding 50,000 (*3*). Understanding the landscape of tree space is important because heuristic algorithms for inferring trees using maximum likelihood (ML), maximum parsimony (MP), and Bayesian inference navigate through parts of this space guided by notions of its structure [e.g., (*4*)]. Moreover, analyses that use phylogenies to study evolutionary processes typically sample from tree space to obtain a good statistical "prior" distribution of phylogenetic relationships used in subsequent comparative analyses, but the design of sampling strategies hinges on the structure of tree space (*5*).

An important advance in understanding tree space was the formulation of the concept of "islands" of trees with similar MP or ML optimality scores (*6*, *7*). Trees belong to the same island if they are near each other in tree space and have optimality scores of L or better with respect to some data matrix. Distance in tree space can be measured by the number of rearrangements required to convert one tree to another. Nearest neighbor interchanges (NNIs), for example, are rearrangements obtained by swapping two subtrees around an internal branch of a tree. Conflicting signals or missing data can result in multiple large tree islands, separated by "seas" of lower-scoring trees, a landscape that can only be characterized by lengthy searches through tree space [e.g., (*8*)]. Empirical studies of phylogenetic tree islands flourished in the context of the single-locus data sets that were common in the 1990s. However, maintaining the same level of accuracy in the larger trees studied today requires combining multiple loci (*9*). The most widely used protocol for data combination is concatenation of multiple alignments of orthologous sequences, one next to another, analyzed as one "supermatrix," a procedure justified when gene tree discordance is low between loci (*10*). Notably, a hallmark of almost all large supermatrix studies is a sizable proportion of missing entries.

[1]Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, AZ 85721, USA. [2]School of Plant Sciences, University of Arizona, Tucson, AZ 85721, USA. [3]Allan Wilson Centre for Molecular Ecology and Evolution, University of Canterbury, Christchurch, New Zealand.

*To whom correspondence should be addressed. E-mail: sanderm@email.arizona.edu