CONCEPTS & SYNTHESIS
EMPHASIZING NEW IDEAS TO STIMULATE RESEARCH IN ECOLOGY

Ecology, 88(8), 2007, pp. 1877–1888

A GLOBAL EVALUATION OF METABOLIC THEORY AS AN EXPLANATION FOR TERRESTRIAL SPECIES RICHNESS GRADIENTS

Bradford A. Hawkins,[1,18] Fabio S. Albuquerque,[2] Miguel B. Araújo,[3,4] Jan Beck,[5] Luis Mauricio Bini,[6] Francisco J. Cabrero-Sañudo,[7] Isabel Castro-Parga,[8] José Alexandre Felizola Diniz-Filho,[6] Dolores Ferrer-Castán,[9] Richard Field,[10] José F. Gómez,[3] Joaquín Hortal,[3,4] Jeremy T. Kerr,[11] Ian J. Kitching,[12] Jorge L. León-Cortés,[13] Jorge M. Lobo,[3] Daniel Montoya,[2] Juan Carlos Moreno,[8] Miguel Á. Olalla-Tárraga,[2] Juli G. Pausas,[14] Hong Qian,[15] Carsten Rahbek,[4] Miguel Á. Rodríguez,[2] Nathan J. Sanders,[16] and Paul Williams[17]

[1] *Department of Ecology and Evolutionary Biology, University of California, Irvine, California 92697 USA*
[2] *Departamento de Ecología, Universidad de Alcalá, 28871 Alcalá de Henares, Madrid, Spain*
[3] *Departamento de Biodiversidad y Biología Evolutiva, Museo Nacional de Ciencias Naturales (CSIC), 28006 Madrid, Spain*
[4] *Center for Macroecology, Institute of Biology, University of Copenhagen, DK-2100 Copenhagen, Denmark*
[5] *Department of Environmental Sciences, Institute of Biogeography, University of Basel, CH-4056, Basel, Switzerland*
[6] *Departamento de Biologia Geral, ICB, Universidade Federal de Goiás, CP 131, 74.001-970, Goiânia, GO, Brazil*
[7] *Departamento de Biodiversidad y Ecología Animal, Instituto de Ecología A.C., A.P. 63, Km 2,5 Ctra. antigua a Coatepec 351, Cong. El Haya, 91070 Xalapa, Veracruz, Mexico*
[8] *Departamento de Biologia, C/ Darwin 2, Universidad Autónoma de Madrid, 28049 Madrid, Spain*
[9] *Área de Ecología, Facultad de Biología, Universidad de Salamanca, 37007 Salamanca, Spain*
[10] *School of Geography, University of Nottingham NG7 2RD United Kingdom*
[11] *Department of Biology, University of Ottawa, Ottawa, Ontario K1N 6N5 Canada*
[12] *Department of Entomology, The Natural History Museum, Cromwell Road, London SW7 5BD United Kingdom*
[13] *Departamento de Ecología y Sistemática Terrestre, El Colegio de la Frontera Sur, Carr. Panamericana y Av. Periférico Sur S/N, San Cristóbal de las Casas, Chiapas 29290 Mexico*
[14] *Fundación Centro de Estudios Ambientales del Mediterráneo (CEAM), C/ Charles R. Darwin 14, Parc Tecnologic, 46980 Paterna, Valencia, Spain*
[15] *Research and Collections Center, Illinois State Museum, 1011 East Ash Street, Springfield, Illinois 62703 USA*
[16] *Department of Ecology and Evolutionary Biology, University of Tennessee, Knoxville, Tennessee 37996 USA*
[17] *The Natural History Museum, Biogeography and Conservation Laboratory, Cromwell Road, London SW7 5BD United Kingdom*

Abstract. We compiled 46 broadscale data sets of species richness for a wide range of terrestrial plant, invertebrate, and ectothermic vertebrate groups in all parts of the world to test the ability of metabolic theory to account for observed diversity gradients. The theory makes two related predictions: (1) ln-transformed richness is linearly associated with a linear, inverse transformation of annual temperature, and (2) the slope of the relationship is near −0.65. Of the 46 data sets, 14 had no significant relationship; of the remaining 32, nine were linear, meeting prediction 1. Model I (ordinary least squares, OLS) and model II (reduced major axis, RMA) regressions then tested the linear slopes against prediction 2. In the 23 data sets having nonlinear relationships between richness and temperature, split-line regression divided the data into linear components, and regressions were done on each component to test prediction 2 for subsets of the data. Of the 46 data sets analyzed in their entirety using OLS regression, one was consistent with metabolic theory (meeting both predictions), and one was possibly consistent. Using RMA regression, no data sets were consistent. Of 67 analyses of prediction 2 using OLS regression on all linear data sets and subsets, two were consistent with the prediction, and four were possibly consistent. Using RMA regression, one was consistent (albeit weakly), and four were possibly consistent. We also found that the relationship between richness and temperature is both taxonomically and geographically conditional, and there is no evidence for a universal response of diversity to temperature. Meta-analyses confirmed significant heterogeneity in slopes among data sets, and the combined slopes across studies were significantly lower than the range of slopes predicted by metabolic theory based on both OLS and RMA regressions. We conclude that metabolic theory, as currently formulated, is a poor predictor of observed diversity gradients in most terrestrial systems.

Key words: diversity gradients; ectotherm diversity; enzyme kinetics; invertebrate diversity; latitudinal gradient; metabolic theory of ecology; plant diversity; species richness; temperature gradients; terrestrial species; vertebrate diversity.

Manuscript received 31 August 2006; accepted 27 October 2006. Corresponding Editor: A. M. Ellison.
[18] E-mail: bhawkins@uci.edu

INTRODUCTION

Although hypotheses to explain biogeographic-scale gradients in species richness date from the 18th and 19th centuries (Forster 1778, von Humboldt 1808), in the past quarter century a number of explanations that include climatic, areal, evolutionary, or "random" components have attracted attention (e.g., Wright 1983, Rohde 1992, Huston 1994, Rosenzweig 1995, Colwell and Lees 2000, Wiens and Donoghue 2004). Recently, the still developing "metabolic theory of ecology" (MTE) has been proposed as an explanation of a range of macroecological patterns, including diversity gradients, by linking ecological and evolutionary processes to plants' and animals' metabolic rates (Allen et al. 2002, 2007, Brown et al. 2004). Although controversial at various levels (e.g., Cyr and Walker 2004, Koehl and Wolcott 2004, Sterner 2004, Whitfield 2004, Muller-Landau et al. 2006*a*, *b*, van der Meer 2006), this theory differs from most previous ones by making fairly precise predictions about the relationship between broadscale patterns of species richness and the proposed environmental driving variable, namely, temperature. In principle, these predictions make the theory testable and falsifiable.

Empirical evaluations of how well observed richness patterns fit the central predictions of MTE are now appearing in the literature (Allen et al. 2002, Kaspari et al. 2004, Hunt et al. 2005, Roy et al. 2006, Algar et al. 2007), although to date they have been taxonomically or geographically limited. As proponents have argued that MTE accounts for diversity gradients over a range of spatial scales from mountain slopes to continental and global gradients, and for many groups of plants and ectothermic animals (Allen et al. 2002, Brown et al. 2004), it is necessary to test the generality of the theory's predictions to evaluate its robustness. In this paper we use a large selection of richness data sets to evaluate MTE. Our focus is on "broadscale" patterns (ranging from hundreds of kilometers to global in extent), and our database includes all suitable richness data for terrestrial plants and ectothermic animals of which we are aware and for which the data were available. We have excluded freshwater and marine systems, although some groups that spend part of their life cycle in freshwater are included (i.e., amphibians; see Allen et al. 2002). Our goal is to test the two primary predictions of MTE as it relates to species richness patterns: (1) ln-transformed species richness is linearly associated with an inverse rescaling of annual temperature, and (2) the slope of the relationship is constrained to be near a particular value specified by the theory. Although testing prediction 1 is reasonably straightforward, testing prediction 2 is complicated by changes in the presentation of the theory as it has evolved.

In the version of the theory described by Allen et al. (2002), annual temperature was rescaled using the transformation 1000/K, where K is kelvins. Allen et al. (2002) claimed that the slope of the relationship between ln-transformed richness and 1000/K should be −9.0. Brown et al. (2004) and Allen et al. (2007) subsequently used a different rescaling, $1/(k\mathrm{K})$, where k is the Boltzman's constant [0.0000862], and claimed a predicted slope of about −0.65. However, these latter papers did not explain clearly that the version of MTE presented by Allen et al. (2002) assumed an energy of activation of 0.78 eV, whereas Brown et al. (2004) and Allen et al. (2007) used a value of 0.65 eV ($1\ \mathrm{eV} = 1.602\ 18 \times 10^{-19}$ J). It is also unclear from the papers why the energy of activation was modified. Irrespectively, this alteration caused Algar et al. (2007) to test North American richness data for six plant and animal groups against an out-of-date prediction, because although they updated the temperature transformation, they did not use the revised activation energy. Similarly, tests of MTE using the version in Allen et al. (2002) (see, e.g., Kaspari et al. 2004, Hunt et al. 2005, Roy et al. 2006) are affected by the change in the energy of activation, as the predicted slope of −9.0 reported in Allen et al. (2002) becomes −7.5 using the more recent activation energy estimate. Uncertainty about the value(s) of activation energies continues (Brown et al. 2003, Enquist et al. 2003).

A second complicating factor is that Allen et al. (2002) used Model II reduced major axis (RMA) regression to test observed slopes of richness–temperature relationships, whereas subsequent analyses used Model I ordinary least squares (OLS) regression on some of the same data sets (Brown et al. 2004, Allen et al. 2007). Because the slope of a RMA regression is equal to the slope of an OLS regression divided by the correlation coefficient of the two variables, the approaches generate different slopes when the correlation between richness and temperature is not 1.0, and so the choice of regression method has serious implications for accepting or rejecting the theory when evaluating real data sets. Despite this uncertainty, most recent tests have used model I regression, under the assumption that temperature data are likely to contain much less error than diversity estimates. In this paper we test the version of MTE presented by Brown et al. (2004) using the same apparent methodological and statistical protocols, although we use both Model I OLS and Model II RMA regression to evaluate the potential effects of the statistical method on our conclusions.

A third level of complexity that arises when testing the predictions of MTE is that the underlying assumptions of the theory have been ignored, even in the original papers that claim to show support for it (Allen et al. 2002, Brown et al. 2004; see also Hunt et al. 2005, Roy et al. 2006). Allen et al. (2002) show mathematically that the energetic-equivalence rule can be used to predict changes in the diversity of ectotherms along temperature gradients only when abundance and average body mass are held constant across samples or communities. Allen et al. (2002) go on to say that these assumptions are supported by the Gentry tropical tree database analyzed by Enquist and Niklas (2001), but they also claim that

MTE is relatively robust when these assumptions are not met. Consistent with their view of the model's robustness, Allen et al. (2002) then show the relationship between the inverse of temperature and the natural log of richness in North American trees in cells generated in a mixed grid of 2.5° × 2.5° in the south and 2.5° × 5.0° in the north, trees along an elevational gradient in Costa Rica, North American amphibians in 2.5° × 2.5° and 2.5° × 5.0° grid cells, Ecuadorian amphibians and Costa Rican amphibians along elevational gradients, fish in watersheds around the world, and prosobranch gastropods in latitudinal bands. In none of these data sets is abundance (whether measured by the number of individuals or density) or body mass held constant. The results for North American trees (erroneously referred to as amphibians) and Costa Rican amphibians were reproduced in Brown et al. (2004), and the North American tree and amphibian results were again presented in Allen et al. (2007). Other tests also have not controlled either variable (Kaspari et al. 2004, Hunt et al. 2005, Roy et al. 2006). Thus, either MTE is robust enough to ignore the assumptions that abundance and body mass must be constant, as the proponents suggest in their papers, or it is not. This has never been formally tested, but in this paper we follow the lead of Allen et al. (2002) and subsequent papers with respect to the types of data that can be used to evaluate MTE, relaxing the assumptions that abundance and body mass must be invariable. This is essential when using geographically extensive data of large grain, as abundance estimates are impossible to obtain over hundreds to hundreds of thousands of square kilometers, and body masses can be estimated crudely at best.

Methods

The data

We analyzed 46 published and unpublished data sets of species richness and annual temperature for a wide range of plants, invertebrates, and ectothermic vertebrates (Tables 1 and 2 and Supplement). All data sets are of medium-to-large geographical extent, the smallest covering Catalonia in northeastern Spain and the largest being global. Most richness estimates were generated using grids with cell sizes (grains) ranging from 100 km^2 to 611 000 km^2. Forty data sets use equal-area or nearly equal-area grids, whereas in five (Californian plants and butterflies, and Chinese plants, reptiles and amphibians) data were aggregated in political units or reserves, so grain size is more variable. We use large-grain data both because of their availability and because proponents consider such data appropriate for evaluating MTE (Allen et al. 2002, 2007, Brown et al. 2004). However, one data set (New World ants) consists of local-scale richness (using observed values), which we include to compare against the previous evaluation of MTE for ants by Kaspari et al. (2004). The taxonomic breadth of the data sets ranges from tribes (bumble bees) to divisions (vascular plants). Methodological details of how the richness data were generated and sources of the temperature data are provided in the original publications (for published data sets) or in the Appendix (for unpublished data sets).

Analytical protocols

The nature of the predictions of MTE required a multistep evaluation of the data. Following Allen et al. (2002) and Brown et al. (2004), we first excluded cells in all data sets with a richness of 0. We then examined each data set throughout its full range of temperatures to determine if the association with ln-transformed richness was linear, the first prediction of MTE. For each data set, we first fitted an OLS linear regression, followed by a split-line regression using the Nonlinear Estimation module in Statistica (quasi-Newton method) (StatSoft 2003) if there appeared to be any nonlinearity in the data. If the analysis identified a break point in the data, we then tested the slopes using a t test. If the two slopes were not significantly different ($P > 0.05$), the relationship between rescaled temperature and ln-transformed richness was classified as being linear throughout the range of the data, whereas data sets with significantly different slopes were classified as being nonlinear. This method identified 23 cases with no significant nonlinearity, including 14 cases having no significant relationship at all. In contrast, there was significant heterogeneity in the slope in 23 cases, which is sufficient reason in itself to reject the first prediction of MTE for those data, at least as the theory is developed and tested in Allen et al. (2002) and Brown et al. (2004). But it remains possible that the theory could explain richness patterns at geographic extents smaller than those chosen by the workers generating the data sets (i.e., the second prediction could hold for part of the data). To examine this possibility, we used the slopes generated by the split-line regressions to compare the data in the "warmer" regions to the left of the break points and the "cooler" regions to the right of the break points. In four data sets, visible nonlinearity remained even after data were divided into subgroups (Fig. 1A, C: n, n′, o′, t′), but patterns were in the opposite direction than those predicted by MTE, so violating the assumption of linearity did not affect our evaluation of the theory for these cases. For North American reptiles, the relationship between richness and temperature was both smoothly curvilinear and the slope was generally negative (Fig. 1A: f), so our conclusion could be influenced by using an inappropriate regression model. Consequently, this case was scored as failing to meet prediction 1, but was not evaluated with respect to prediction 2. Our analytical method resulted in 67 OLS regressions against which to compare slopes with MTE. This procedure was then repeated using RMA regression (Sokal and Rohlf 1995), generating an additional 67 slopes that were compared against prediction 2.

Because gridded species richness data usually contain small-scale spatial autocorrelation due to the proximity

TABLE 1. Summary of regressions testing Model I (OLS) and Model II (RMA) slopes of richness–temperature relationships for cases with linear relationships between rescaled temperature and ln-transformed richness.

Group	Region	Figure	r^2	P	OLS slope	RMA slope
Blister beetles	North America	c	0.35	0.001	**−0.49**	**−0.83**
Ants	Colorado/Nevada	l	0.05	0.30	+0.34	+1.52
Hawk moths	Mexico	m	0.22	0.002	**−0.84**	**−1.79**
Reptiles	Brazil	o	0.01	0.75	+0.35	+3.50
Tiger beetles	northwestern South America	p	0.16	0.009	**−0.57**	**−1.43**
Ants	New World	q	0.58	0.008	**−0.87**	**−1.14**
Butterflies	Australia	r	0.03	0.51	+0.32	+1.85
Amphibians	Australia	s	<0.01	0.85	−0.08	−0.80
Tiger beetles	Australia	t	0.11	0.08	−0.48	−1.45
Dung beetles	Iberia/France	v	<0.01	0.44	−0.12	−1.20
Reptiles	Europe	w	0.61	0.001	**−0.79**	**−1.01**
Plants (native)	Great Britain	c′	0.52	0.06	−0.14	−0.19
Amphibians	Iberia	e′	0.01	0.42	−0.15	−1.50
Reptiles	Iberia	f′	<0.01	0.68	+0.07	+0.70
Pteridophytes	Iberia	g′	0.06	0.07	+0.50	+2.04
Seed plants	Iberia	h′	0.08	0.01	**+0.38**	**+1.34**
Plants	Catalonia	i′	0.04	0.16	+0.28	+1.40
Orthoptera	Catalonia	j′	0.10	0.008	**+0.77**	**+2.43**
Woody plants	southern Africa	k′	0.02	0.64	−0.41	−2.90
Reptiles	southern Africa	l′	<0.01	0.93	+0.01	+0.10
Tiger beetles	India	p′	<0.01	0.94	+0.02	+0.20
Reptiles	China	q′	0.38	0.002	**−0.61**	**−0.99**
Amphibians	China	r′	0.40	0.002	**−0.53**	**−0.84**

Notes: OLS is ordinary least squares; RMA is reduced major axis. "Figure" letters refer to the panels in Fig. 1A–C in which data sets are illustrated. Also provided are the coefficients of determination for each regression (r^2) and significance levels. Significance tests are based on the geographically effective degrees of freedom (v^*), estimated using the modified t test of Dutilleul (1993), and slopes that are significant at $P < 0.05$ are in bold. See Supplement: Table S1 for standard errors and 95% confidence intervals of slopes, raw sample sizes, geographically effective degrees of freedom, and sources of the richness data.

of grid cells (Legendre and Legendre 1998, Diniz-Filho et al. 2003), model residuals are not independent, which can cause nonsignificant relationships to appear significant. To reduce the Type I error in the regressions, we tested the significance of slopes based on the geographically effective degrees of freedom using the modified t test of Dutilleul (1993), implemented in SAM (spatial analysis in macroecology; Rangel et al. 2006) and derived from spatial correlograms of both variables. However, this test is conservative because it assumes that all spatial autocorrelation is artifactual, so to reduce the resulting Type II error we truncated the correlogram to only the positive Moran's I values to correct the degrees of freedom. We did not adjust significance levels for the large number of tests (i.e., Bonferroni correction), as this would also generate excessive Type II error (Gotelli and Ellison 2004).

After generating slopes for each data set, we tested for a common combined slope following the meta-analytical approach described by Zeka et al. (2003) and Zeka and Schwartz (2004). First, a weighted-average slope $\bar{b}_{\mathrm{w}}$ was calculated using the reciprocals of the squared standard errors ($1/\mathrm{SE}^2$) of the slopes as weights (Hillebrand et al. 2001). To take the spatial autocorrelation within data sets into account, these standard errors were first corrected by the geographically effective degrees of freedom. The homogeneity statistic Q (following a χ^2 distribution) was also used to test the homogeneity of the slopes across all studies. Rejection of the hypothesis of homogeneity implies that the effect should be considered random and the weighted-average slope cannot be considered an estimate of the common slope (with all studies sharing a single slope). Instead, a grand-mean slope can be calculated by using weights that assume random variation among the studies. The grand-mean slope and associated standard error were also estimated (see equations in Zeka et al. [2003]) and were then used to test more formally the slope predicted by MTE. Meta-analyses were performed separately on OLS and RMA slopes.

Criteria for acceptance of MTE

Brown et al. (2004) argued that slopes of richness–temperature regressions should fall between −0.60 and −0.70, although they interpreted observed slopes "close" to this range as also providing support for the theory. To reflect this uncertainty, we evaluated data fit as follows: (1) nonlinear relationships do not support the claim of MTE that taxa respond to temperature identically everywhere; (2) clearly nonsignificant relationships (defined as $P > 0.10$) do not support the theory that temperature drives the diversity gradients of plants and ectothermic animals within the range of the data; (3) statistically significant ($P < 0.05$) negative slopes between −0.60 and −0.70 are fully consistent with

Table 2. Summary of regressions testing Model I (OLS) and Model II (RMA) slopes of richness–temperature relationships for cases with nonlinear relationships between rescaled temperature and ln-transformed richness.

				Cool				Warm			
Group	Region	Figure	Break point	r^2	Prob.	OLS slope	RMA slope	r^2	Prob.	OLS slope	RMA slope
Bumble bees	global	a	41.5	0.11	0.06	−0.23	−0.69	0.48	0.02	**+0.79**	**+1.14**
Snakes	Afrotropics	n′	38.2	0.05	0.14	−0.35	−1.57	0.33	<0.001	**+2.80**	**+4.87**
Amphibians	Afrotropics	o′	38.2	0.05	0.22	−0.55	−2.46	0.38	<0.001	**+3.25**	**+5.27**
Woody plants	Kenya	m′	39.0	0.03	0.32	+0.20	+1.15	0.05	0.02	**+0.08**	**+0.36**
Eupelmid wasps	Palearctic	b′	40.8	0.79	<0.001	**−0.67**	**−0.75**	0.31	0.03	**+0.32**	**+0.57**
Butterflies	western Palearctic	a′	41.1	0.38	0.07	−0.57	−0.92	0.62	0.008	**+1.25**	**+1.59**
Dung beetles	western Palearctic	z	40.8	0.68	0.002	**−0.39**	**−0.47**	0.52	<0.001	**+0.46**	**+0.64**
Pteridophytes	Europe	u	41.3	0.01	0.56	−0.07	−0.70	0.18	0.06	+1.08	+2.55
Amphibians	Europe	x	41.2	0.56	0.03	**−1.07**	**−1.43**	<0.01	0.99	0	0
Trees	Europe	y	41.2	0.48	0.06	−0.73	−1.05	0.01	0.67	−0.08	−0.80
Plants (exotic)	Great Britain	d′	41.2	0.42	<0.001	**−4.76**	**−7.34**	0.29	<0.01	**−3.14**	**−5.83**
Trees	North America	d	42.1	0.59	<0.001	**−1.06**	**−1.38**	0.13	0.24	−0.33	−0.92
Butterflies (w)	North America	g	42.2	0.23	0.06	−0.35	−0.73	<0.01	0.55	−0.01	−0.10
Butterflies (s)	North America	h	42.2	0.25	0.04	**−0.37**	**−0.74**	0.11	0.007	**−0.09**	**−0.27**
Tiger beetles	North America	b	41.6	0.57	0.001	**−1.04**	**−1.38**	0.01	0.68	−0.05	−0.50
Amphibians	North America	e	42.1	0.72	<0.001	**−0.88**	**−1.04**	0.36	0.06	−0.48	−0.80
Grasshoppers	North America	i	41.5	0.28	0.07	−0.57	−1.08	0.13	0.15	+0.25	+0.69
Reptiles	North America	f	NA								
Plants	California	j	40.7	0.24	0.003	**−0.34**	**−0.69**	0.48	<0.001	**+0.71**	**+1.02**
Butterflies	California	k	40.3	0.15	0.02	**+0.15**	**+0.39**	0.11	<0.001	**+0.25**	**+0.75**
Amphibians	Brazil	n	39.1	0.30	0.03	**+0.52**	**+0.95**	0.38	0.006	**+1.58**	**+2.56**
Hawk moths	Southeast Asia	t′	39.4	0.19	0.21	−0.07	−0.16	0.29	0.23	+0.44	+0.82
Angiosperms	China	s′	40.7	0.46	0.003	**−0.54**	**−0.80**	0.02	0.50	+0.12	+0.85

Notes: Break point is the rescaled temperature at which the relationship changes slope within each data set (see Fig. 1A–C). Cases for the parts of the data to the right of each break point (Cool) are listed first, followed by the parts of the data to the left of each break point (Warm). Columns are as defined in Table 1. North American butterflies are distinguished by winter (w) and summer (s) distributions. The reptiles of North America could not be analyzed using split-line regression (NA, not applicable).

the theory as presented by Brown et al. (2004); and (4) marginally significant ($0.05 < P < 0.10$) slopes or slopes between −0.55 and −0.59 or −0.71 and −0.75 could possibly be consistent with the theory.

Although we calculated 95% CIs for all slopes (see Supplement: Table S1), we do not use the usual evaluation of model fit (by conducting t tests of the predicted slope against observed slopes) for two reasons. First, proponents of the most current versions of MTE accept a range of slopes rather than a precise slope as representing reasonable fits. Second, the standard approach invites Type II error with respect to rejecting MTE, because the weaker the relationship between temperature and richness, the wider the standard error of the slope and the more difficult it is to reject the theory. To circumvent this problem, the combined slopes from the meta-analyses were compared to the range of predicted slopes (−0.60 to −0.70) to evaluate overall congruence of observed slopes with MTE. This was done for OLS and RMA separately.

RESULTS

Linear data sets

Twenty-three data sets had approximately linear responses of richness to temperature (i.e., no significant heterogeneity in slopes throughout the range of the data). However, 14 of these had no significant relationship at all (Table 1), allowing us to reject the first prediction of MTE for these cases. These latter data sets are distributed widely around the Earth, although most are found in regions with warm climates. Of the remaining nine cases with significant richness–temperature relationships, slopes were negative in seven, but only one (Chinese reptiles, Fig. 1C: q′) fell within the range of slopes predicted by MTE when analyzed using OLS regression. Thus, we reject prediction 2 of MTE in 22 of 23 cases. No cases were within the predicted range using RMA regression. Relaxing the statistical level of significance of the regression to $P = 0.10$ and expanding the acceptable range of slopes to −0.55 through −0.75 generated possible agreement with the theory for tiger beetles in northwestern South America (Fig. 1A: p) using OLS, although the r^2 of this regression was 0.16, indicating that temperature is a very poor predictor of tiger beetle richness irrespective of the statistical significance and slope of the relationship. No cases were possibly consistent with prediction 2 using RMA regression.

Nonlinear data sets

Although 23 data sets had nonlinear relationships with temperature, which is inconsistent with prediction 1, it remains possible that prediction 2 could be supported in at least parts of the data. Indeed, in 10 cases the slope was significantly negative in the cooler parts of the data (Table 2). However, only the small family of parasitic wasps Eupelmidae within part of the western Palearctic (Fig. 1B: b′, data to the right of the

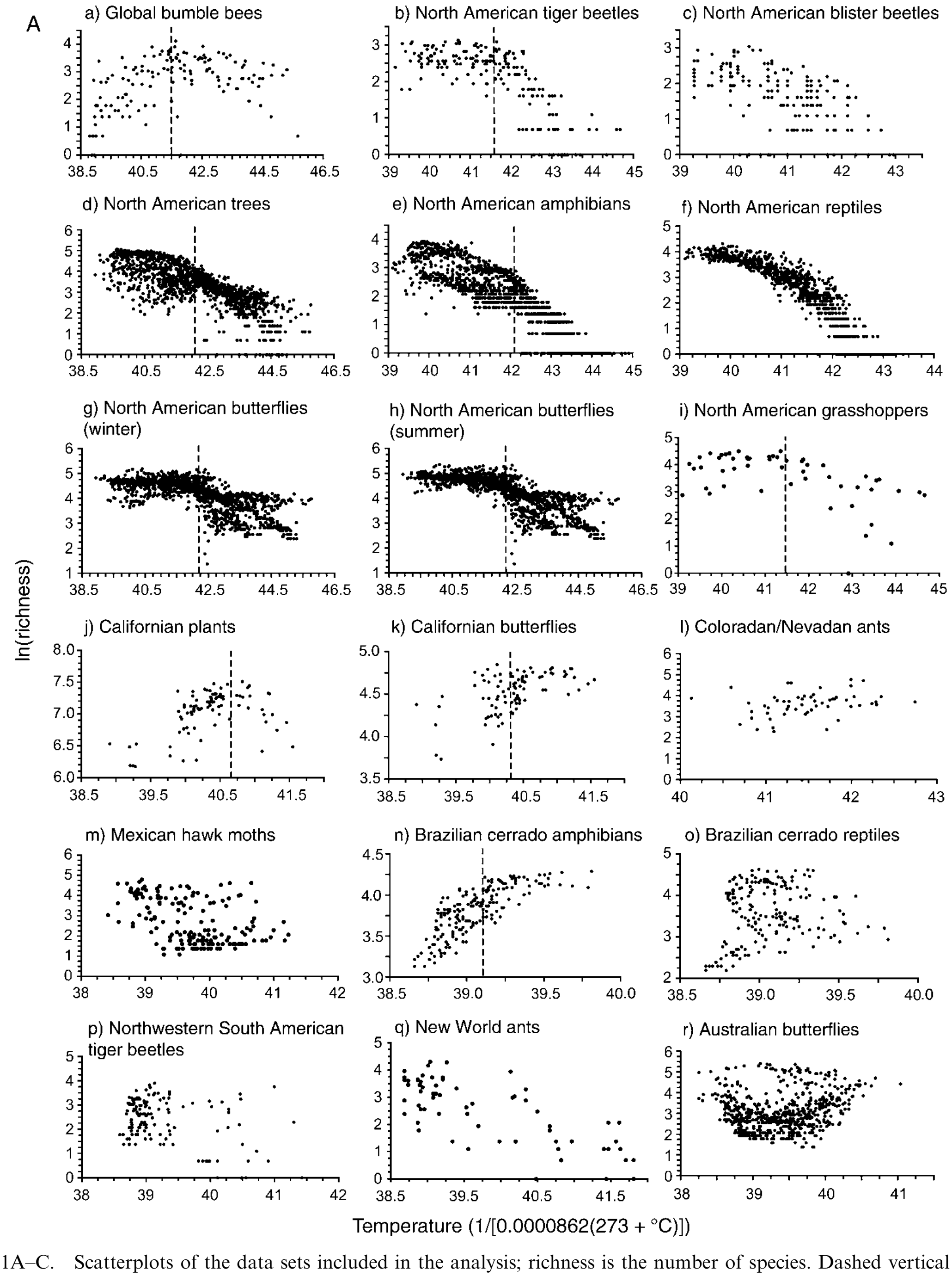

FIG. 1A–C. Scatterplots of the data sets included in the analysis; richness is the number of species. Dashed vertical lines are break points used to divide nonlinear data into linear components. Note that the temperature variable is a reciprocal; actual temperature decreases from left to right. "Fig. 1A" refers to the panels on this page; Figs. "1B" and "1C" are on the following pages.

break point) was fully consistent with the theory based on the slope of OLS regressions, whereas only Californian plants conformed using RMA (but with a low coefficient of determination [0.24]). In 10 cases, richness had no statistically significant relationship with temperature. Expanding both the range of acceptable slopes and the significance level added Canadian grasshoppers (Fig. 1A: i, data to the right of the break point) and northern European trees (Fig. 1B: y, data to the right of the break point) and butterflies (Fig. 1B: a′, data to the right of the break point) as possibly conforming to the theory using OLS. Using RMA, Canadian butterflies (Fig. 1A: g, data to the right of the break point) became congruent, and four data sets were possibly congruent:

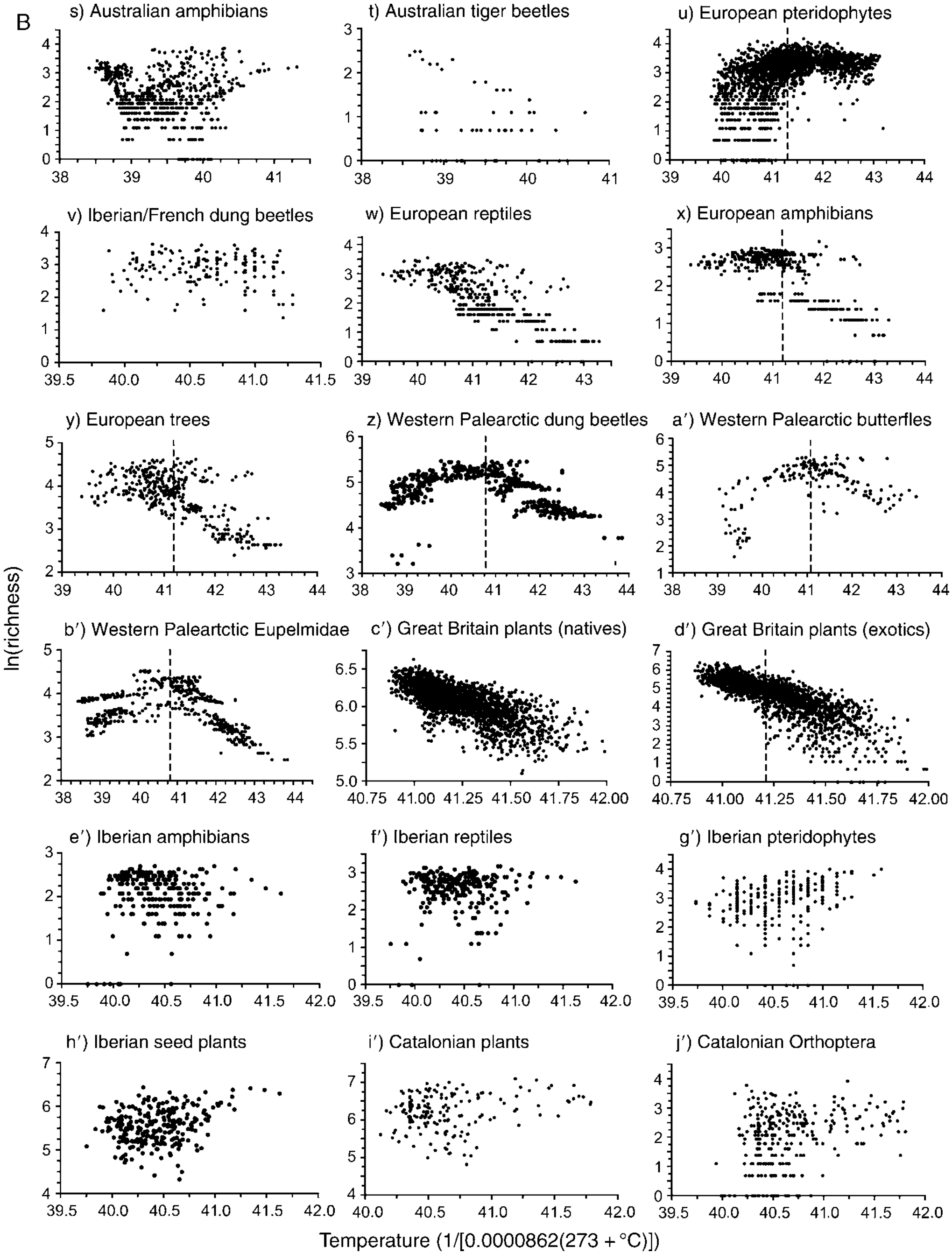

FIG. 1B. Continued.

bumble bees in the northern temperate zone (Fig. 1A: a, data to the right of the break point), Canadian butterflies in both summer and winter (Fig. 1A: g–h, data to the right of the break point), and European eupelmids, which changed from fully confirmatory using OLS to possibly confirmatory). Irrespectively, even in parts of the world with cold or very cold climates, results clearly consistent with the theory were found in only one of 22 cases, whether using OLS or RMA regressions, and the specific cases differed depending on the method of analysis.

In the warmer parts of these data sets, richness was either independent of rescaled temperature or decreased with increasing rescaled temperature (the slope was positive) in 20 of 22 cases (Table 2). In the two cases with significant negative slopes, both were far from the prediction. These results are similar for both OLS and RMA slopes.

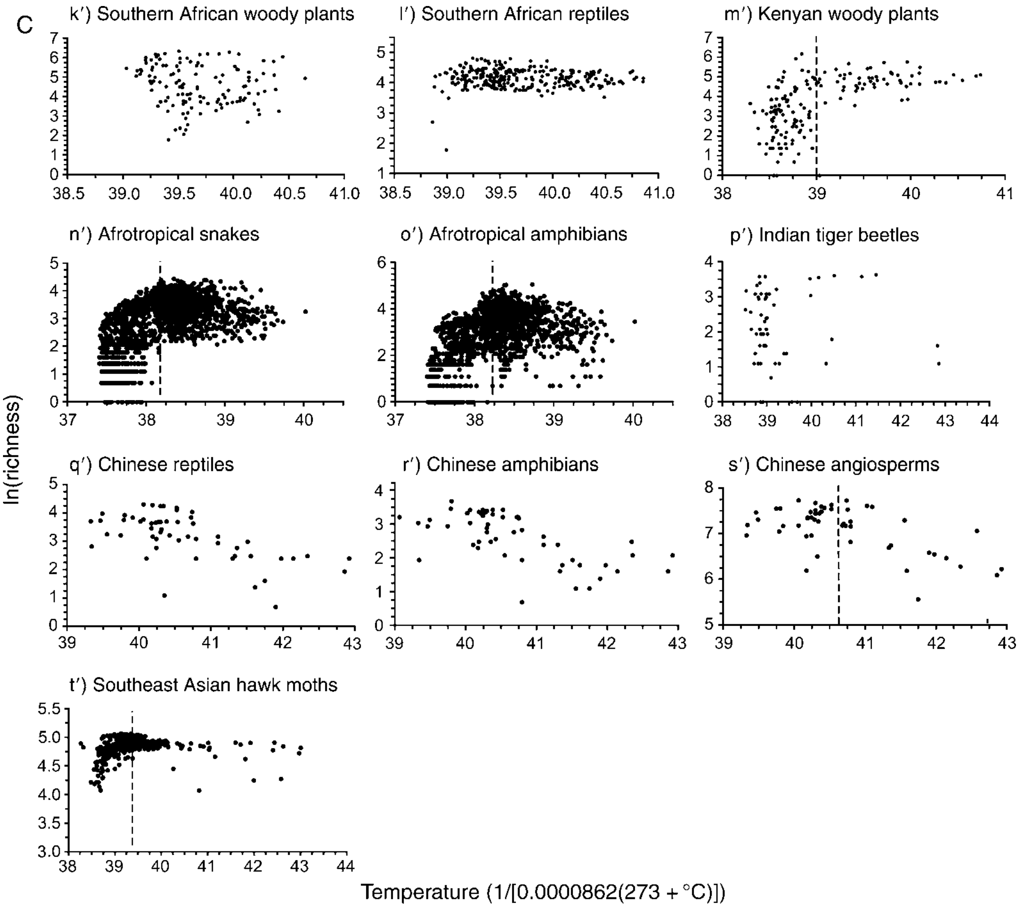

FIG. 1C. Continued.

The data for the reptiles of North America did not allow split-line regression, as the relationship with temperature is smoothly nonlinear (Fig. 1A: f). Therefore we could not evaluate prediction 2, other than to note that species richness generally increases with temperature, but with no consistent slope.

Overall, based on 67 OLS regressions, the predictions of MTE were supported in two (one weakly due to a low r^2), and four additional cases were possibly consistent with the theory. RMA regression generated a similar level of support for MTE (one fully consistent case and four possible cases).

Meta-analysis

The combined slopes across all studies, estimated by the OLS and RMA regressions, were +0.085 and −0.255, respectively. These are far from the predicted range of slopes put forth by Brown et al. (2004), and the OLS slope is positive rather than negative. Further, the homogeneity statistics indicated that the variation among studies is highly significant in both cases (Q = 9108.5 and 524.6; both $P < 0.001$). The grand-mean slopes ± SE ("means of the distributions of study slopes" [Zeka et al. 2003]), were −0.105 ± 0.0005 (OLS regression) and −0.143 ± 0.0016 (RMA regression), both substantially shallower than predicted by MTE.

DISCUSSION

Our most salient result is that we find very limited support for the supposition that the metabolic theory of ecology accurately predicts geographic diversity gradients, as the theory has been described in Allen et al. (2002) and Brown et al. (2004). Using strict statistical conditions for evaluation of data results in only one or two regressions consistent with the theory, depending on whether OLS or RMA regression is considered the appropriate method of analysis. It is possible that a few additional data sets could be consistent with the theory if the criteria for acceptance are broadened, but this does not alter the conclusion that the prediction presented in Brown et al. (2004) is not supported across a wide range of terrestrial plants and ectothermic animals with data gathered at regional, continental, and intercontinental extents. Pooling data sets for meta-analysis does not alter this conclusion.

Because MTE uses a specific aspect of climate (temperature) to explain richness gradients, and climate is believed to have stronger effects on richness at larger extents (Hawkins et al. 2003, Whittaker et al. 2003), it could be argued that the spatial extents of many of the data sets that we examined are too small to represent "fair" tests of the theory. However, pragmatically, the geographic scale of richness patterns does not influence our evaluation of MTE, because few data sets are consistent with the theory, whatever their extent or grain. Further, the single most consistent case (with an OLS slope close to the prediction [–0.67] and a high coefficient of determination [0.79]) is restricted to northern and central Europe, which is not a geographically extensive region; even for this case the prediction breaks down when the extent is increased to include the western Palearctic. The single apparently confirmatory case using RMA covers only the cooler parts of California, also not a large region, and the regression has a low coefficient of variation. But irrespective of the extent of individual data sets, currently existing data of continental, hemispherical, and global extents most commonly show that richness is not linearly associated with temperature or other measures of energy at very large extents (Fig. 1A–C; see also Hawkins et al. 2003, Algar et al. 2007, Whittaker et al. 2007), further eroding support for metabolic theory as a general explanation for observed global-scale diversity gradients.

We also find that temperature does not explain large amounts of variance in richness gradients in most parts of the world, which would be unexpected if the mechanism underlying MTE were the primary driving force of diversity everywhere. Annual mean temperature explains $<50\%$ of the variance in richness in 56 (84%) of our regressions, and $<25\%$ in 38 (57%); see Table 1. Perhaps most surprising is that in the 33 regressions in which rescaled temperature and richness are significantly correlated, the slope is positive in 14 (recalling that MTE uses a reciprocal of temperature as the predictor, so a positive slope means that richness decreases with increasing temperature). It is notable that negative slopes (i.e., positive correlations between actual temperature and richness) are concentrated in data sets that extend into areas with cold winters, whereas in warm temperate, subtropical, and tropical data sets, richness is independent of temperature or decreases with increasing temperature. This is consistent with the well-known importance of water to terrestrial richness gradients in both plants and animals, especially in warm climates (e.g., Rahbek and Graves 2001, Hawkins et al. 2003). However, the key importance of water to diversity is not restricted to hot places, at least for plants. For example, Hawkins et al. (2007) found that the richness gradient of North American and European trees is better explained statistically by rainfall than by annual temperature ($r^2 = 0.706$ vs. 0.525, respectively), and, unlike the case with temperature, the association of tree richness with rainfall is linear across all climates. Based on an overwhelming amount of evidence for the importance of water to life, any explanation for diversity gradients that depends solely on temperature will probably be incomplete (see also Field et al. 2005, O'Brien 2006).

Given that richness is usually negatively correlated with inverse temperature in northern latitudes, the key question becomes: how close to the slope predicted by MTE must an observed slope be to be consistent with the theory? Clearly, a negative slope per se is not strong support, because all proposed mechanisms for the "latitudinal" diversity gradient are intended to explain why there are fewer species at the poles. The presumed strength of MTE over most of the alternative theories is that it makes a fairly precise prediction, which should make it falsifiable. But, as that prediction has to be made increasingly vague to encompass the range of observed patterns, the more difficult it becomes to differentiate MTE from alternative explanations. The problem of precision is potentially serious; initial formulations of MTE assumed a precise energy of activation (Allen et al. 2002), whereas latter papers have admitted that they may be highly variable (Brown et al. 2003, Enquist et al. 2003), leading to a wide range of slopes (Brown et al. 2003). But if virtually any negative slope is considered evidence for an influence of enzyme kinetics on richness, richness–temperature relationships cannot be used to distinguish MTE from other explanations that predict more species in the tropics, and indeed the model becomes untestable in any situation in which richness increases with temperature. Even so, when richness *decreases* with increasing temperature, as occurs in many parts of the world, MTE can be rejected without equivocation unless energies of activation are allowed to take biologically impossible values.

A related point in testing MTE arises from the assumptions of the model as formulated by Allen et al. (2002), especially the temperature invariance of density and body mass. Although we tested the predictions of the model as presented in the original papers by Allen et al. (2002) and Brown et al. (2004), it is potentially important to evaluate how violations in underlying assumptions will affect the shape of the relationship between richness and temperature. It is possible that knowing how density and body size vary along temperature gradients will allow a more generalized version of the model of Allen et al. (2002), although this will strongly restrict the types of data that can be used to evaluate MTE. Based on the equations in Allen et al. (2002), if density decreases or average body mass changes along a temperature gradient, we expect nonlinear relationships between temperature and richness, but a detailed investigation of the theoretical aspects of the model are beyond the scope of this paper. However, these issues may be critical for understanding how violating the assumptions of body size and density invariance affects predictions of the model.

A fourth aspect of our results relates to claims that MTE applies to many taxonomic groups over most or all parts of the planet. If it were true that enzyme kinetics were the primary influence on diversity patterns,

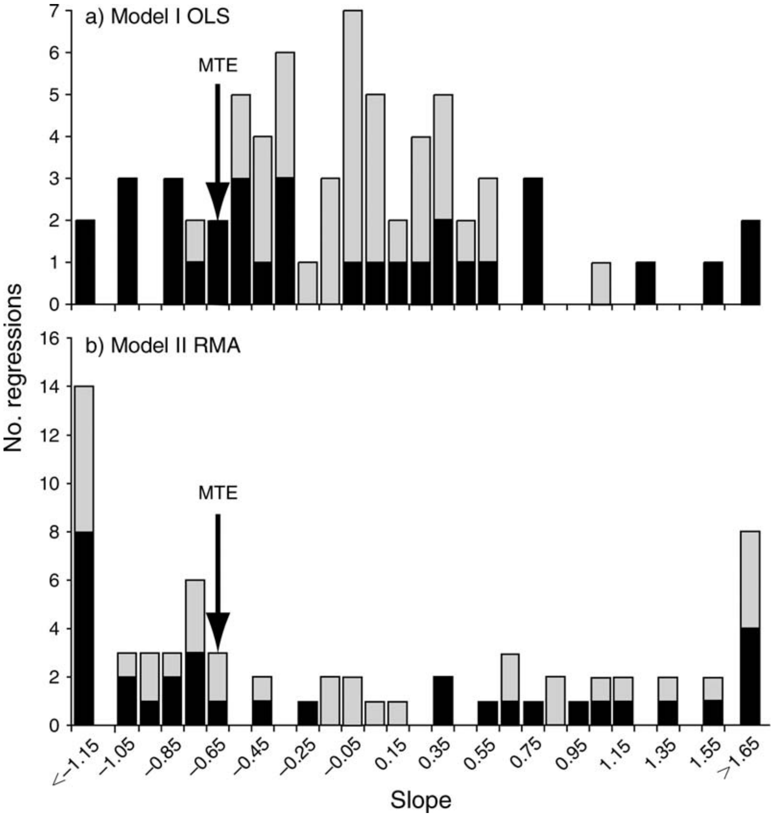

FIG. 2. Frequency distribution of slopes of all (a) OLS and (b) RMA regressions (see Table 1). The arrows identify the class containing the slope predicted by metabolic theory. Black bars represent statistically significant ($P < 0.05$) regressions, whereas nonsignificant regressions ($P > 0.05$) are in gray.

we would at least expect observed slopes of inverse temperature–richness regressions to cluster around the predicted slope of −0.65, even if they did not all have exactly that slope due to variable activation energies and secondary influences on diversity that might be taxonomically or geographically specific (see Brown et al. 2003). However, when we plot the distribution of slopes from the linear regressions, there is no tendency for slopes to be distributed around the predicted value, irrespective of the regression method used or whether or not they are statistically significant (Fig. 2). More importantly, the meta-analytical results were clear-cut, with grand-mean slopes much lower than −0.65 (in both OLS and RMA regression analyses). Based on the variability in the relationships between temperature and species richness across the studies (as indicated by the highly significant heterogeneity of slopes over studies), we must conclude that the responses of plants and animals to temperature are both taxonomically and geographically conditioned and, consequently, there is no universal explanation for diversity gradients driven by the speed of chemical reactions across all temperatures and taxa. It does not follow that temperature does not influence diversity gradients, but it seems unlikely that MTE can be the primary force driving diversity patterns in terrestrial systems at the extents represented in our data sets (from hundreds of kilometers to global). This will be the case even if future studies find examples in which slopes are consistent with the theory.

It also does not appear that heterogeneity in responses of organisms to temperature is related to their general biology, or that plants, invertebrates, and ectothermic vertebrates differ fundamentally in their response. The slopes of neither OLS nor RMA regressions differ significantly among the three groups (one-way ANOVAs; for OLS, $F_{2,64} = 1.70$, $P = 0.190$; for RMA, $F_{2,64} = 1.13$, $P = 0.328$). The relationships of richness with temperature depend much more strongly on where the organisms occur than on what taxonomic group is being considered (see Fig. 1A–C). This is expected, because most groups in our collection of data sets have broadly congruent diversity patterns, being least diverse in deserts and polar climates and most diverse in the warm, wet tropics.

We are unable to duplicate previous results claimed to be consistent with metabolic theory (Brown et al. 2004, Kaspari et al. 2004). In the cases of North American trees and amphibians as reported by Allen et al. (2002, 2007) and Brown et al. (2004), their conclusion depended on fitting linear regressions through curvilinear data (see also Algar et al. 2007) and, thus, the presumed support derived from averaging slopes that are too steep in the north and too shallow in the south. Because interpreting linear regression coefficients when applied to curvilinear relationships is questionable, we believe that the conclusions in Allen et al. (2002) and Brown et al. (2004) are not compelling. In the case of ants as reported by Kaspari et al. (2004), the discrepancy arises solely because they tested the version of the theory that assumed an energy of activation of −0.78 (see *Introduction*). After correcting the prediction of the formulation of Allen et al. (2002) with the new activation energy, the observed OLS slope of Kaspari et al. is much shallower than the −7.5 slope predicted by MTE ($b = -2.8$), whereas their RMA slope is too steep ($b = -9.0$). The meta-analysis of Hunt et al. (2005) similarly can be reevaluated. Across 10 fossil foraminiferan data sets, they found an average RMA slope of −10.7, which is substantially steeper than the MTE prediction when using the more recent energy of activation. Further, this average includes one data set with a slope of +24.0, and when this strongly inconsistent relationship was excluded, the mean decreased to −14.8 (95% CL: −18.9 and −10.7) (Hunt et al. 2005:742), significantly too steep to conform to MTE. On the other hand, the subsequent analysis of Roy et al. (2006) for seven data sets of marine gastropods and bivalves reported three slopes close to −7.5 and four with slopes ranging between −2.7 and −5.8. All studies taken together suggest that previous support for MTE was not as strong as may have been assumed; consequently, based on previous analyses as well as our own, there is currently little empirical support for claims that MTE accurately predicts diversity gradients.

The unresolved issue of whether richness–temperature relationships should be tested using Model I or Model II regression does not affect our general conclusion. Our evaluation of individual data sets and the meta-analyses generate very similar results whether we use OLS or RMA approaches. Thus, the method of analysis makes no practical difference to our conclusion that data rarely fit the theory. On the other hand, this does matter with respect to determining whether any particular data set is consistent with MTE or not when the temperature–richness correlation is not strong. When tests of the theory are applied to individual cases, serious attention must be paid to determining which regression method is most appropriate for the data.

In sum, although we cannot conclude that enzyme kinetics have no role to play in explaining broadscale patterns of diversity, we can conclude that there is very limited supporting evidence that observed richness gradients are consistent with the predictions of MTE, in its current form, across a wide range of taxonomic groups in almost all regions of the world. It is important to stress that we have restricted our evaluation of MTE to diversity gradients and have tested a specific form of the theory. We also use data that some might argue are inappropriate (the data sets may contain variable body sizes and abundances across the temperature gradients), although these criticisms also apply to data that were claimed to support the theory. Further, we cannot directly evaluate the theory's ability to explain aquatic diversity gradients. It is obvious that in terrestrial systems water is essential for any diversity at all, and it is possible that in systems where water is not limiting, enzyme kinetics could explain observed gradients. Finally, our focus has been on "broadscale" diversity gradients. Smaller scale gradients, such as those along mountain slopes, might also conform better to MTE predictions. Future analyses can address these possibilities.

ACKNOWLEDGMENTS

We thank three anonymous reviewers for their careful dissection of the first version of the manuscript. J. Hortal is supported by a Portuguese FCT grant (BPD/20809/2004), M. A. Rodríguez was supported by the Spanish CICYT (grant REN2003-03989/GLO), and J. A. F. Diniz-Filho and L. M. Bini are supported by productivity grants from Brazilian CNPq.

LITERATURE CITED

Algar, A. C., J. T. Kerr, and D. J. Currie. 2007. A test of metabolic theory as the mechanism underlying broad-scale species-richness gradients. Global Ecology and Biogeography 16:170–178 [doi: 10.1111/j.1466-8238.2006.oo275x].

Allen, A. P., J. F. Gillooly, and J. H. Brown. 2002. Global biodiversity, biochemical kinetics, and the energetic-equivalence rule. Science 297:1545–1548.

Allen, A. P., J. F. Gillooly, and J. H. Brown. 2007. Recasting the species–energy hypothesis: the different roles of kinetic and potential energy in regulating biodiversity. Pages 283–299 *in* D. Storch, P. A. Marquet, and J. H. Brown, editors. Scaling biodiversity. Cambridge University Press, Cambridge, UK.

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

Brown, J. H., J. F. Gillooly, G. B. West, and V. M. Savage. 2003. The next step in macroecology: from general empirical patterns to universal ecological laws. Pages 408–423 *in* T. M. Blackburn and K. J. Gaston, editors. Macroecology: concepts and consequences. Blackwell Publishing, Oxford, UK.

Colwell, R. K., and D. C. Lees. 2000. The mid-domain effect: geometric constraints on the geography of species richness. Trends in Ecology and Evolution 15:70–76.

Cyr, H., and S. C. Walker. 2004. An illusion of mechanistic understanding. Ecology 85:1802–1804.

Diniz-Filho, J. A. F., L. M. Bini, and B. A. Hawkins. 2003. Spatial autocorrelation and red herrings in geographical ecology. Global Ecology and Biogeography 12:53–64.

Dutilleul, P. 1993. Modifying the *t*-test for assessing the correlation between two spatial processes. Biometrics 49: 305–314.

Enquist, B. J., E. P. Economo, T. E. Huxman, A. P. Allen, D. D. Ignace, and J. F. Gillooly. 2003. Scaling metabolism from organisms to ecosystems. Nature 423:639–942.

Enquist, B. J., and K. J. Niklas. 2001. Invariant scaling relations across tree-dominated communities. Nature 410: 655–660.

Field, R., E. M. O'Brien, and R. J. Whittaker. 2005. Global models for predicting woody plant richness from climate: development and evaluation. Ecology 86:2263–2277.

Forster, J. R. 1778. Observations made during a voyage round the world, on physical geography, natural history, and ethnic philosophy. G. Robinson, London, UK.

Gotelli, N., and A. M. Ellison. 2004. A primer of ecological statistics. Sinauer Associates, Sunderland, Massachusetts, USA.

Hawkins, B. A., R. Field, H. V. Cornell, D. J. Currie, J. F. Guégan, D. M. Kaufman, J. T. Kerr, G. G. Mittelbach, T. Oberdorff, E. M. O'Brien, E. E. Porter, and J. R. G. Turner. 2003. Energy, water, and broad-scale geographic patterns of species richness. Ecology 84:3105–3117.

Hawkins, B. A., D. Montoya, M. Á. Rodríguez, M. Á. Olalla-Tárraga, and M. Á. Zavala. 2007. Global models for predicting woody plant richness from climate: comment. Ecology 88:255–259.

Hillebrand, H., F. Watermann, R. Karez, and U.-G. Berninger. 2001. Differences in species richness patterns between unicellular and multicellular organisms. Oecologia 126:114–124.

Hunt, G., T. M. Cronin, and K. Roy. 2005. Species–energy relationship in the deep sea: a test using the Quaternary fossil record. Ecology Letters 8:739–747.

Huston, M. A. 1994. Biological diversity. The coexistence of species on changing landscapes. Cambridge University Press, Cambridge, UK.

Kaspari, M., P. S. Ward, and M. Yuan. 2004. Energy gradients and the geographic distribution of local ant diversity. Oecologia 140:407–413.

Koehl, M. A. R., and B. D. Wolcott. 2004. Can function at the organismal level explain ecological patterns? Ecology 85: 1808–1810.

Legendre, P., and L. Legendre. 1998. Numerical ecology. Elsevier, Amsterdam, The Netherlands.

Muller-Landau, H. C., et al. 2006*a*. Testing metabolic ecology theory for allometric scaling of tree size, growth and mortality in tropical forests. Ecology Letters 9:575–588.

Muller-Landau, H. C., et al. 2006*b*. Comparing tropical forest tree size distributions with the predictions of metabolic ecology and equilibrium models. Ecology Letters 9:589–602.

O'Brien, E. M. 2006. Biological relativity to water–energy dynamics. Journal of Biogeography 33:1868–1888.

Rahbek, C., and G. R. Graves. 2001. Multiple scale assessment of patterns of avian species richness. Proceedings of the National Academy of Sciences (USA) 98:4534–4539.

Rangel, T. F. L. V. B., J. A. F. Diniz-Filho, and L. M. Bini. 2006. Towards an integrated computational tool for spatial analysis in macroecology and biogeography. Global Ecology and Biogeography 15:321–327.

Rohde, K. 1992. Latitudinal gradients in species diversity: the search for the primary cause. Oikos 65:514–527.

Rosenzweig, M. L. 1995. Species diversity in space and time. Cambridge University Press, Cambridge, UK.

Roy, K., D. Jablonski, and J. W. Valentine. 2006. Beyond species richness: biogeographic patterns and biodiversity dynamics using other metrics of diversity. Pages 151–170 *in* M. V. Lomolino and L. R. Heaney, editors. Frontiers of biogeography: new directions in the geography of nature. Sinauer Associates, Sunderland, Massachusetts, USA.

Sokal, R. R., and F. J. Rohlf. 1995. Biometry: the principles and practice of statistics in biological research. Third edition. Freeman, New York, New York, USA.

StatSoft. 2003. STATISTICA, Version 6. ⟨http://www.statsoft.com⟩

Sterner, R. W. 2004. A one-resource "stoichiometry"? Ecology 85:1813–1815.

van der Meer, J. 2006. Metabolic theories in ecology. Trends in Ecology and Evolution 21:136–140.

von Humboldt, A. 1808. Ansichten der Natur mit wissenschaftlichen Erläuterungen. J. G. Cotta, Tübingen, Germany.

Whitfield, J. 2004. Ecology's big, hot idea. PloS Biology 2(12): e440 [doi: 10.1371/journal.pbio.0020440].

Whittaker, R. J., D. Nogués-Bravo, and M. B. Araújo. 2007. Geographic gradients of species richness: a test of the water–energy conjecture of Hawkins et al. (2003*a*) using European data for five taxa. Global Ecology and Biogeography 16:76–89.

Whittaker, R. J., K. J. Willis, and R. Field. 2003. Climatic–energetic explanations of diversity: a macroscopic perspective. Pages 107–129 *in* T. M. Blackburn and K. J. Gaston, editors. Macroecology: concepts and consequences. Cambridge University Press, Cambridge, UK.

Wiens, J. J., and M. J. Donoghue. 2004. Historical biogeography, ecology and species richness. Trends in Ecology and Evolution 19:639–644.

Wright, D. H. 1983. Species–energy theory: an extension of species–area theory. Oikos 41:496–506.

Zeka, A., R. Gore, and D. Kriebel. 2003. Effects of alcohol and tobacco on aerodigestive cancer risks: meta-regression analysis. Cancer Causes and Control 14:897–906.

Zeka, A., and J. Schwartz. 2004. Estimating the independent effects of multiple pollutants in the presence of measurement error: an application of a measurement-error-resistant technique. Environmental Health Perspectives 112:1686–1690.

APPENDIX

Methods for data sets not available in the literature (*Ecological Archives* E088-112-A1).

SUPPLEMENT

Summary regression statistics and sources for all data sets (*Ecological Archives* E088-112-S1).

Metabolic Theory and Patterns of Species Richness[1]

Identifying and explaining biogeographic gradients in species richness is a long-standing preoccupation of ecologists. By and large, terrestrial species richness increases toward the equator and decreases with elevation, but most explanations of these patterns are phenomenological; they originate from statistical correlations whose significance is based on a trivial test of the null hypothesis that there is no relationship between richness and the hypothesized "causal" variable. The recent axiomatic derivation of the Metabolic Theory of Ecology (MTE) provides a notable exception: based on biochemical kinetics of metabolism, MTE presents a precise, quantitative prediction of how species richness should vary with (appropriately scaled) environmental temperature: linearly, with slope ≈ -0.65 (J. P. Allen, J. H. Brown, and J. F. Gillooly. 2002. "Global biodiversity, biochemical kinetics, and the energetic-equivalence rule." Science 297:1545–1548).

In the focal Concepts and Synthesis paper for this Forum, Hawkins et al. test the predictions of MTE with 46 different data sets compiled for a variety of terrestrial plants, invertebrates, and ectothermic vertebrates; the geographic extents of these data sets range from a small region in Spain to continents and hemispheres. How might this confrontation between theory and data have come out? First, richness–temperature relationships all could be linear with slopes whose confidence intervals include –0.65. Such a result would be a powerful confirmation of MTE. Second, a few of the data sets might not fit the predictions. Such a result would support MTE and would highlight interesting situations in which ecological or evolutionary constraints might alter fundamental scaling relationships. For example, water availability or nutrient limitation may interact strongly with temperature in water-limited systems, leading to unexpected relationships between temperature and richness (Brown, J. H., J. F. Gillooly, A. P. Allen, V. M. Savage, and G. B. West. 2004. "Toward a metabolic theory of ecology." Ecology 85:1771–1789). Last, most data sets could fail to support MTE. This is the result that Hawkins et al. find, and they conclude that MTE is a poor predictor of observed biogeographic patterns of species richness. Whether this result is due to shortcomings of MTE, inappropriate data, or different methods of statistical analyses is explored in further detail by Latimer and by Gillooly and Allen in their commentaries on Hawkins et al.'s paper.

My goal in assembling this Forum was first and foremost to provide a jumping-off point for future studies of MTE and species richness. Progress in this field will be most rapid when a theory whose structure is agreed upon is tested with standard methods and suitable data sets. These include: the correct scaling of temperature; a fixed and clearly defined activation energy of metabolism; whether to use ordinary least-squares (OLS) or reduced-major-axis (RMA) regression, and even whether a linear relationship ought to be assumed; and the appropriate taxonomic ranges and geographic scales of the data. It is critically important that all participants in this debate agree on core axioms and assumptions.

How we test theories depends on their structure and presentation. Brown et al. (2004: 1787) asserted that "[a] strength of [MTE] is that it makes explicit quantitative predictions based on first principles." Allen et al. (2002), Brown et al. (2004), and now Hawkins et al. use a strong, deductive approach to test MTE: Is the observed relationship between temperature and richness the same as that predicted by the theory? An inductive approach, in which theory is continually refined as more observations accumulate, provides opportunities for synthesis from which general theories may eventually emerge. Such an approach is cautiously advocated by Gillooly and Allen. Bayesian inference allows one to formalize induction. There is thus a certain irony that the hierarchical Bayesian reanalysis of Hawkins et al.'s data by Latimer fails to provide additional support for MTE as a mechanism underlying latitudinal gradients in species richness, *given the data at hand.*

A new theory generates much excitement, but this initial excitement can be tempered as the theory is scrutinized carefully and challenged by data. Synthesis emerges from the ongoing confrontation of theory and data, but its rate of emergence depends on the flexibility and open-mindedness of the participants. By sharing data and offering constructive critiques, the participants in this Forum not only have established the benchmark for future studies of processes underlying biogeographic patterns of species richness; they also have provided an example of how scientific theories evolve and develop.

—Aaron M. Ellison
Associate Editor-in-Chief

Key words: Bayesian inference; deduction; induction; latitudinal gradient; metabolic theory of ecology; species richness.

---

[1] Reprints of this 26-page Forum (including the Hawkins et al. Concepts and Synthesis paper) are available for $10.00 each, either as PDF files or as hard copy. Prepayment is required. Order reprints from the Ecological Society of America, Attention: Reprint Department, 1707 H Street, N.W., Suite 400, Washington, DC 20006, USA.

Ecology, 88(8), 2007, pp. 1890–1894

LINKING GLOBAL PATTERNS IN BIODIVERSITY TO EVOLUTIONARY DYNAMICS USING METABOLIC THEORY

James F. Gillooly[1,3] and Andrew P. Allen[2]

[1]*Department of Zoology, 223 Bartram Hall, University of Florida, Gainesville, Florida 32611 USA*
[2]*National Center for Ecological Analysis and Synthesis, 735 State St., Santa Barbara, California 93101 USA*

Introduction

Starting in 2002, with a paper entitled "Global biodiversity, biochemical kinetics and the energetic-equivalence rule," we have been developing a theoretical framework to understand the mechanisms underlying broadscale biodiversity gradients, particularly the latitudinal gradient. This work is part of a broader Metabolic Theory of Ecology (MTE) being developed to predict various aspects of the structure and function of ecological systems (Brown et al. 2004). Although MTE has been criticized (see Hawkins et al. 2007), support for its predictions continues to grow (Anderson et al. 2006, Anfodillo et al. 2006, Lopez-Urrutia et al. 2006, Meehan 2006, Robinson 2006). In the preceding paper, Hawkins et al. (2007) criticize the original work of Allen et al. (2002) based on their analyses of a large number of empirical data sets. Here we respond to their major criticisms and discuss important issues raised by their paper.

What Is the Theory as First Proposed by Allen et al. (2002)?

Hawkins et al. (2007) characterize our theory as based only on "the effects of temperature on enzyme kinetics." This oversimplification stems from an incomplete reading of Allen et al. (2002) and subsequent publications. Since Allen et al. (2002), we have published several papers that clarify our position (Allen et al. 2003, Brown et al. 2003), test model assumptions (Savage et al. 2004, Gillooly et al. 2005*b*), and expand on the theory by presenting new derivations and data (Allen and Gillooly 2006, Allen et al. 2006; Allen and Savage, *in press*). Yet, Hawkins et al. (2007) evaluate only a single prediction from Allen et al. (2002). They consider the more "detailed theoretical aspects" of this work, while potentially critical, they admit, as beyond the scope of their paper.

So what is the theory? Our theory aims to uncover the mechanisms controlling the origin and maintenance of biodiversity gradients based on the constraints of energetics on speciation–extinction dynamics. The theory specifies that the process of speciation is influenced by the effects of individual-level variables (i.e., body size, temperature) on rates of genetic divergence among populations, and by the effects of ecosystem-level variables (e.g., net primary production) on the numbers of genetically diverging populations maintained in communities. The process of extinction, as with Hubbell's Neutral Biodiversity Theory, NTB (Hubbell 2001), is assumed to be a function of both the speciation rate and population abundance (Allen et al. 2007; Allen and Savage, *in press*).

Unlike most recent work on species–energy theory, we have proposed that two forms of energy, kinetic energy and chemical potential energy, both help to regulate biodiversity through their effects on rates of speciation and extinction, but in different ways (Fig. 1). Specifically, we have proposed that kinetic energy influences biodiversity through its effects on individual metabolic rate, because metabolic rate constrains evolutionary rates through its effects on rates of individual turnover (1/generation time) and mutation (Allen et al. 2002, 2007). This assumption is now supported by data showing that (1) rates of individual turnover and rates of mutation show the same temperature dependence as metabolic rate (Savage et al. 2004, Gillooly et al. 2005*b*, Allen et al. 2006); (2) rates of speciation in one group of oceanic plankton also show this same temperature dependence (Allen et al. 2006); and finally, that (3) species richness is positively correlated with speciation rates (Emerson and Colm 2005, Allen and Gillooly 2006). These findings are consistent with the "evolutionary speed" hypothesis of Rohde (1978, 1992).

With respect to chemical potential energy, we have proposed that NPP and the factors that control its availability (i.e., water, nutrients, temperature) influence speciation rates through their effects on total community abundance, and therefore the total numbers of genetically diverging populations (Allen et al. 2002, 2006, 2007). This is consistent with the "more individuals hypothesis" (Wright 1983), as well as with NBT (Hubbell 2001).

Thus, Allen et al. (2002) and subsequent papers are developing a quantitative theoretical framework that links ecological and evolutionary dynamics of individ-

---

Manuscript received 20 November 2006; revised 22 December 2006; accepted 22 December 2006. Corresponding Editor: A. M. Ellison.

[3] E-mail: gillooly@zoo.ufl.edu

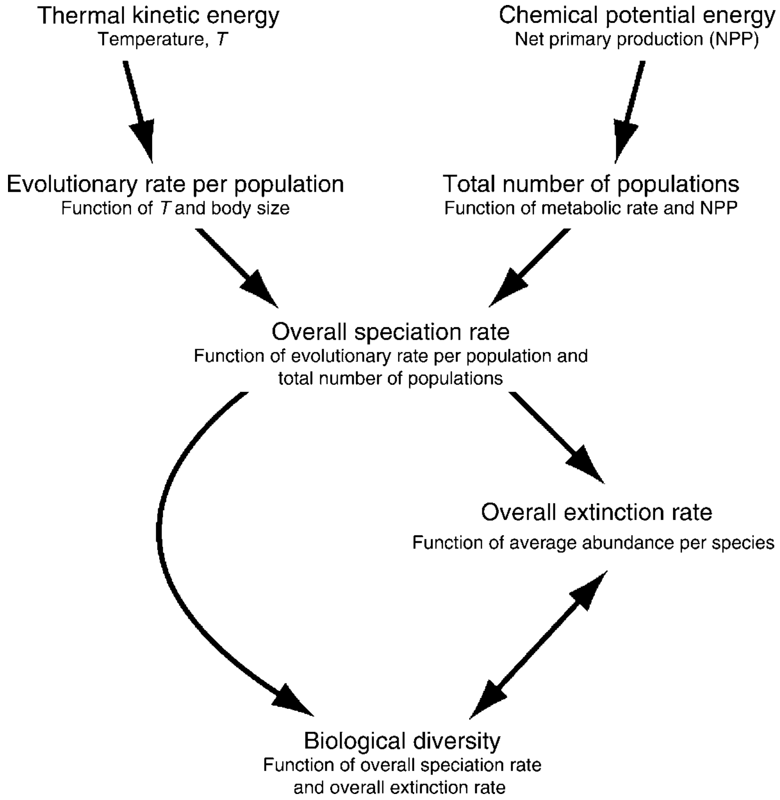

FIG. 1. The theoretical framework proposed by Allen et al. (2007) for the relationship of biological diversity to kinetic and potential forms of energy in the environment (modified from Allen et al. [2007]).

uals and populations to patterns of biodiversity in communities.

WHAT DOES THE THEORY PREDICT?

Contrary to the assertion by Hawkins et al. (2007), Allen et al. (2002) clearly state that "we do not mean to imply that temperature is the only variable that affects biodiversity." We have specifically mentioned that nutrient and water availability are important to consider, and that they may interact strongly in water-limited systems (Brown et al. 2003, 2004, Allen et al. 2005, 2007). Thus, we agree with the central conclusion of Hawkins et al. (2007) that "based on an overwhelming amount of evidence ... any explanation that depends solely on temperature will likely be incomplete." Indeed, our theory predicts that water limitation should constrain biodiversity through its effects on community abundance.

Allen et al. (2002) make several predictions about global-scale gradients in biodiversity and the underlying mechanisms. First, they predict that evolutionary rates should show the same body size and temperature dependence as mass-specific metabolic rate, B, defined as

$$B = b_o M^{-1/4} e^{-E/kT} \tag{1}$$

where b_o is a normalization constant that varies by taxonomic group, M is body mass, E is the average activation energy of metabolic rate (~0.6–0.7 eV, where 1 electronvolt = 1.602×10^{-19} J), k is Boltzmann's constant (8.62×10^{-5} eV/K, where K is degrees kelvin), and T is average body temperature in degrees kelvin. Second, using Eq. 1, Allen et al. (2002) extend the energetic-equivalence rule (EER) of Damuth (1987) to predict the combined effects of size and temperature on community abundance. Extensive population abundance data for endotherms and ectotherms were presented in support of this prediction. These data indicated that the total energy flux per population per unit area, B_T, was independent of body size and body temperature.

Based on this EER, Allen et al. (2002) then predict that species richness, S, in plots of fixed area, A, should be described by the following function for both endotherms and ectotherms:

$$S = (J/A)(B_o/B_T)e^{-E/kT}. \tag{2}$$

In this expression, B_o (the normalization constant for metabolic rate) is a function of organism size and b_o, B_T varies by taxonomic group and with plot area A, and J/A is the total density of individuals per unit area (see Allen et al. 2002). Note that because Eq. 2 includes temperature and community abundance, it encompasses both the "evolutionary speed" hypothesis and the "more individuals" hypothesis. In other words, this expression attempts to unify ecological and evolutionary explanations for species richness.

As a result, Eq. 2 yields different predictions for endotherms vs. ectotherms. For endotherms, Eq. 2 predicts that species richness, S, should increase linearly with total abundance per unit area, J/A, irrespective of environmental temperature, provided that the size distribution of organisms is held constant. This prediction is consistent with a recent global-scale analysis of bird community data (Pautasso and Gaston 2005). Conversely for ectotherms, species richness should vary as a function of abundance, body size, and environmental temperature. So, when abundance and size are both held constant, Eq. 2 predicts that plots of $\ln(S)$ vs. $1/kT$ (inverse temperature) should yield a straight line with a slope of ~ -0.6 to -0.7 eV. Hawkins et al. (2007) focus on evaluating only this latter prediction.

Do Data Support the Theory?

Allen et al. (2002) evaluated the predicted temperature dependence of species richness using global scale data on richness for both aquatic and terrestrial taxonomic groups (trees, amphibians, fish, fish ectoparasites, and gastropods). In doing so, they were careful to point out that the assumptions of size and abundance invariance with respect to temperature gradients must be viewed as "working hypotheses." Given this caveat, data from Allen et al. (2002) and other researchers (e.g., Kaspari et al. 2004), are broadly supportive of this model prediction.

In contrast, Hawkins et al. (2007) test the theory and conclude that there is virtually no support for this particular prediction of Allen et al. (2002) based on data from terrestrial ecosystems. However, their conclusion depends strongly on their choice of data and methodology, as well as on their interpretation of the data. The following are a few examples.

First, nearly half of the data sets presented by Hawkins et al. (2007) are for restricted groups of insects (e.g., Eupelmidae wasps). Allen et al. (2002:1546) are careful to point out that model assumptions are "not expected to hold true for groups that are narrowly defined...." Total insect diversity clearly peaks in the warm tropics, but restricted taxa can show markedly different patterns (e.g., Ichneumonid wasps; Janzen 1981). Furthermore, and importantly, many of these insect groups maintain relatively constant body temperatures in different thermal environments, including the bumble bees, sphinx moths, dung beetles, and butterflies (Heinrich 1981; see also Bartholomew and Heinrich [1978] for dung beetles, Heinrich and Vogt [1993] for bumble bees, Heinrich and Casey [1973] for moths). Thus, these groups would not be predicted to show the same exponential temperature dependence for species richness as shown by true ectotherms.

Second, the authors include data sets in which extreme gradients in water availability occur in the opposite direction of temperature (e.g., the African data sets, which include the Sahara), and they exclude data sets for which water availability is not an issue (i.e., aquatic ecosystems), on the basis that such data are beyond the scope of their paper. In cases of extreme water limitation, they show that richness often decreases with increasing temperature and consider this to be evidence that temperature is not operating in the same way across systems. This is not necessarily the case. For example, Hawkins et al. (2007) show that plant richness in Catalonia decreases with increasing temperature. Yet, the original publication by these authors concludes that, after correcting for the effects of water availability and other variables, richness actually increases exponentially with temperature (Pausas et al. 2003: Fig. 2). In other words, these results are consistent with the prediction of Allen et al. (2002).

Third, Hawkins et al. (2007) do not consider the quality and extent of the data sets that they have assembled, and how these differences might affect their results. A review of the original publications shows that many of the data sets in Hawkins et al. (2007) measure richness, and especially temperature, in different ways. For example, in their measure of richness, the study of Amazonian amphibians did not include any species that could not be identified or that were part of a "species group" (Diniz-Filho et al. 2006; data from Diniz-Filho et al. [2004]). In the case of the Australian tiger beetles, the very limited data consist of only 1–3 species in many areas, which led the original authors to conclude "the small sample sizes make *any* interpretation questionable" (Pearson and Juliano 1993:201). In measuring temperature, the authors point to the methods of the original publications, but many of these studies did not measure temperature (e.g., Field et al. 2005), or used measures of temperatures that were quite different (e.g., highest temperature of the decade in Hawkins and Porter [2003]). In short, the haphazard compilation of data by Hawkins et al. (2007) raises serious questions about the applicability of their analyses and the conclusions that they have reached.

Still, in spite of these problems, Hawkins et al. (2007) raise some important and interesting questions about how the theory of Allen et al. (2002) should be confronted with data. Hawkins et al. (2007) use a strict Popperian approach that aims to falsify the theory based on best-fit statistical criteria of a single-model prediction. In doing so, they find that 42% of the data sets that show linear relationships with temperature reject the "null" prediction (i.e., 0.6–0.7 eV activation energy) of Allen et al. (2002), based on the 95% confidence intervals. But, they also argue that data sets exhibiting "significant" nonlinearity can be viewed as showing no support for the prediction of Allen et al. (2002). Based on these criteria, they break nonlinear data sets in two and report many weaker relationships for lines fit through only a portion of the data.

From our perspective, this approach by Hawkins et al. (2007) results in unreasonably casting aside this

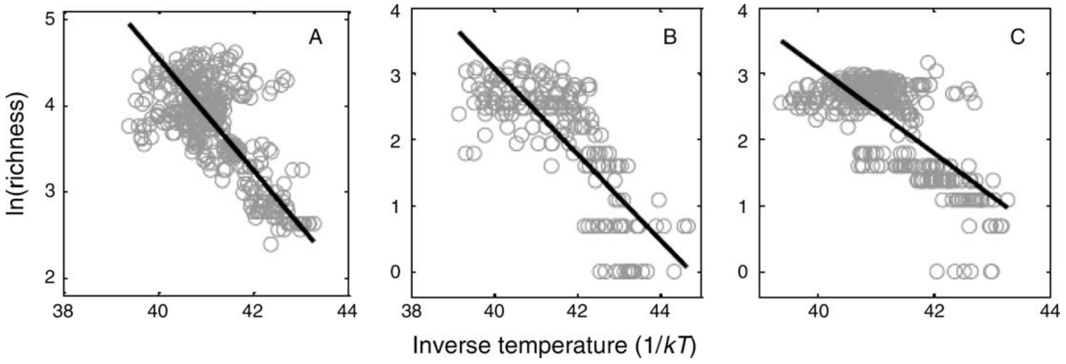

FIG. 2. Species richness–temperature relationships that Hawkins et al. (2007) conclude are not supportive of the Allen et al. (2002) model. Here, k is Boltzmann's constant (8.62×10^{-5} eV/K, where K is degree kelvin), and T is average body temperature in degrees kelvin. The straight lines in the figure, with a slope of −0.65, represent the relationship predicted by Allen et al. (2002). Results are for: (A) European trees ($r^2 = 0.37$), (B) North American tiger beetles ($r^2 = 0.54$), and (C) European amphibians ($r^2 = 0.50$). Lines are fit using best-fit criteria, where the slope is fixed at −0.65 (r^2 values are based on these fits). Note that we have not plotted the two best examples of this result to avoid reusing the data sets from Allen et al. (2002), but the same result holds even more strongly for these data (North American amphibians, $r^2 = 0.74$; North American trees, $r^2 = 0.58$).

young theory even though its assumptions and predictions are supported by considerable (but not all) data. This approach risks losing the knowledge that could be gained by pursuing the theory further. For example, in Fig. 2, we have included three data sets from Hawkins et al. (2007), chosen because they are among the most supportive of the predictions of Allen et al. (2002), and thus are not intended to be a representative sample. A straight-line fit to these data, with MTE's predicted slope of −0.65 eV, captures the overall relationship of richness to temperature and explains up to 54% of the variation in these relationships. And yet, each of these data sets is significantly nonlinear. Based on this nonlinearity, Hawkins et al. (2007) argue against the model of Allen et al. (2002). From our perspective, these results are promising, and the nonlinearity points to the need to test whether incorporating other variables in the model will improve predictions. In particular, this is because, as Hawkins et al. (2007) mention, violations of the Allen et al. (2002) model assumptions lead to curvilinear relationships in plots of log-transformed richness vs. inverse temperature. The question of which approach or interpretation is better is a philosophical one.

More generally, however, we would argue that the theory of Allen et al. (2002) is best evaluated not just by testing a single prediction, but rather by using a more holistic approach that evaluates all assumptions and predictions of the theory, as well as the logic behind the theory. Ideally, the performance of the model in predicting patterns of biodiversity should also be compared to that of other predictive models.

CAN THE THEORY BE IMPROVED/FURTHER DEVELOPED? IS IT EVOLVING?

Absolutely. Much remains to be done. On the empirical side, testing the predictions of Allen et al. (2002) presents significant challenges. Broadscale data on body size, community abundance, and species diversity are scarce, but they do exist (e.g., Pautasso and Gaston 2005). Both this theory and NTB point to the need for more and better data in these areas. On the theoretical side, the framework of this theory requires a better understanding of how water limitation and other factors control community abundance, and how abundance in turn affects speciation–extinction dynamics. Recent progress on this issue has been made by Allen and Savage (*in press*) by extending NTB. As further insights emerge, or as modifications are required, the theory will continue to evolve. For example, in Allen et al. (2002), the average activation energy, E, was defined as ~0.78eV, based on analyses of Gillooly et al. (2001). Upon further analyses (e.g., Brown et al. 2004, Gillooly et al. 2005*a*, *b*), this value is now consistently defined as 0.6–0.7 eV. Certainly, there is abundant room for additional research on these and other mechanisms.

CLOSING THOUGHTS

Allen et al. (2002) and subsequent papers are developing a promising theoretical framework that links short-term species coexistence to long-term speciation–extinction dynamics. As pointed out by Brown et al. (2003), this research indicates that the effect of temperature on rates of biological metabolism must be *an important component* of any theory that attempts to explain broadscale patterns in biodiversity, such as the latitudinal gradient.

To be sure, this approach differs markedly from the approach that describes patterns in biodiversity using statistical models. Statistical models of this sort typically include a few to several variables with the goal of "explaining" as much variance as possible. The choice of variables or models is admittedly relatively straightforward with these models because they are based solely on best-fit criteria. These models can explain substantial variation in patterns of biodiversity and can identify climatic factors that play an important role. They are less useful in identifying the underlying mechanisms responsible for these patterns, and cannot make a priori predictions.

Both approaches are necessary to build a general theory of biodiversity. We hope that the future will bring more "cross-fertilization" between these and other approaches on this important topic.

ACKNOWLEDGMENTS

We thank Ethan White for his many helpful comments and suggestions.

LITERATURE CITED

Allen, A. P., J. H. Brown, and J. F. Gillooly. 2002. Global biodiversity, biochemical kinetics, and the energetic-equivalence rule. Science 297:1545–1548.

Allen, A. P., J. H. Brown, and J. F. Gillooly. 2003. Response to comment on "Global biodiversity, biochemical kinetics, and the energetic-equivalence rule." Science 299:346.

Allen, A. P., and J. F. Gillooly. 2006. Assessing latitudinal gradients in speciation rates and biodiversity at the global scale. Ecology Letters 9:947–954.

Allen, A. P., J. F. Gillooly, and J. H. Brown. 2005. Linking the global carbon cycle to individual metabolism. Functional Ecology 19:202–213.

Allen, A. P., J. F. Gillooly, and J. H. Brown. 2007. Recasting the species–energy hypothesis: the different roles of kinetic and potential energy in regulating biodiversity. Pages 283–299 *in* D. Storch, P. A. Marquet, and J. H. Brown, editors. Scaling biodiversity. Cambridge University Press, New York, New York, USA.

Allen, A. P., J. F. Gillooly, V. M. Savage, and J. H. Brown. 2006. Kinetic effects of temperature on rates of genetic divergence and speciation. Proceedings of the National Academy of Sciences (USA) 103:9130–9135.

Allen, A. P., and V. M. Savage. *In press*. Setting the absolute tempo of biodiversity dynamics. Ecology Letters.

Anderson, K. J., A. P. Allen, J. F. Gillooly, and J. H. Brown. 2006. Temperature-dependence of biomass accumulation rates during secondary succession. Ecology Letters 9:673–682

Anfodillo, T., V. Carraro, M. Carrer, C. Foir, and S. Rossi. 2006. Convergent tapering of xylem conduits in different woody species. New Phytologist 169:219–221.

Bartholomew, G. A., and B. Heinrich. 1978. Endothermy in African dung beetles during flight, ball making and ball rolling. Journal of Experimental Biology 73:65–83.

Brown, J. H., A. P. Allen, and J. F. Gillooly. 2003. Heat and biodiversity. Response. Science 299:512–513.

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

Damuth, J. 1987. Interspecific allometry of population density in mammals and other animals: the independence of body mass and population energy use. Biological Journal of the Linnaean Society 31:193–246.

Diniz-Filho, J. A. F., L. M. Bini, M. P. Pinto, T. Rangel, P. Carvalho, and R. P. Bastos. 2006. Anuran species richness, complementarity and conservation conflicts in Brazilian cerrado. Acta Oecologica 29:9–15.

Diniz-Filho, J. A. F., L. M. Bini, C. M. Vieira, M. C. de Souza, R. P. Bastos, D. Brandao, and L. G. Oliveira. 2004. Spatial patterns in species richness and priority areas for conservation of anurans in the cerrado region, central Brazil. Amphibia–Reptilia 5:63–75.

Emerson, B. C., and M. Colm. 2005. Species diversity can drive speciation. Nature 434:1015–1017.

Field, R., E. M. O'Brien, and R. J. Whittaker. 2005. Global models for predicting woody plant richness from climate: development and evaluation. Ecology 86:2263–2277.

Gillooly, J. F., A. P. Allen, V. M. Savage, G. B. West, and J. H. Brown. 2005*a*. Response to Clarke and Fraser: Effects of temperature on metabolic rate. Functional Ecology 20:400–404.

Gillooly, J. F., A. P. Allen, G. B. West, and J. H. Brown. 2005*b*. 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.

Gillooly, J. F., J. H. Brown, G. B. West, V. M. Savage, and E. L. Charnov. 2001. Effects of size and temperature on metabolic rate. Science 293:2248–2251.

Hawkins, B. A., et al. 2007. A global evaluation of metabolic theory as an explanation for terrestrial species richness gradients. Ecology 88:1877–1888.

Hawkins, B. A., and E. E. Porter. 2003. Does herbivore diversity depend on plant diversity? The case of California butterflies. American Naturalist 161:40–49.

Heinrich, B. 1981. Insect thermoregulation. John Wiley, New York, New York, USA.

Heinrich, B., and T. Casey. 1973. Metabolic rate and endothermy in sphinx moths. Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology 82:195–206.

Heinrich, B., and F. D. Vogt. 1993. Abdominal temperature regulation by Arctic bumblebees. Physiological Zoology 66:257–269.

Hubbell, S. P. 2001. A unified neutral theory of biodiversity and biogeography. Princeton University Press, Princeton, New Jersey, USA.

Janzen, D. H. 1981. The peak in North American Ichneumonid species richness lies between 38 and 42°N. Ecology 62:532–537.

Kaspari, M., P. S. Ward, and M. Yuan. 2004. Energy gradients and the geographic distribution of local ant diversity. Oecologia 140:407–413.

Lopez-Urrutia, A., E. San Martin, R. P. Harris, and X. Irigoien. 2006. Scaling the metabolic balance of the oceans. Proceedings of the National Academy of Sciences (USA) 103:8739–8744.

Meehan, T. D. 2006. Energy use and animal abundance in litter and soil communities. Ecology 87:1650–1658.

Pausas, J. G., J. Carreras, A. Ferré, and X. Font. 2003. Coarse-scale plant species richness in relation to environmental heterogeneity. Journal of Vegetation Science 14:661–668.

Pautasso, M., and K. J. Gaston. 2005. Resources and global avian assemblage structure in forests. Ecology Letters 8:282–289.

Pearson, D. L., and S. A. Juliano. 1993. Evidence for the influence of historical processes in co-occurrence and diversity of tiger beetle species. Pages 194–202 *in* R. E. Ricklefs and D. Schluter, editors. Species diversity in ecological communities: historical and geographical perspectives. University of Chicago Press, Chicago, Illinois, USA.

Robinson, D. 2006. Biology's big idea. Nature 444:272.

Rohde, K. 1978. Latitudinal gradients in species-diversity and their causes. I. A review of the hypotheses explaining the gradients. Biologisches Zentralblatt 97:393–403.

Rohde, K. 1992. Latitudinal gradients in species diversity: the search for the primary cause. Oikos 65:514–527.

Savage, V. M., J. F. Gillooly, J. H. Brown, G. B. West, and E. L. Charnov. 2004. Effects of body size and temperature on population growth. American Naturalist 163:E429–E441.

Wright, D. H. 1983. Species–energy theory: an extension of species–area theory. Oikos 41:496–506.

Ecology, 88(8), 2007, pp. 1895–1898

GEOGRAPHY AND RESOURCE LIMITATION COMPLICATE METABOLISM-BASED PREDICTIONS OF SPECIES RICHNESS

ANDREW M. LATIMER[1]

Department of Ecology and Evolutionary Biology, University of Connecticut, 75 North Eagleville Road,
Storrs, Connecticut 06269 USA

METABOLIC THEORY OF ECOLOGY: A REANALYSIS OF DATA

Assessments of general theoretical frameworks like neutral ecological theory and the metabolic theory of ecology (MTE) often use only one or two data sets, and thus do not provide much evidence about the particular prediction tested, much less the validity or usefulness of the general theory. By compiling a diverse group of data sets from a range of terrestrial taxa, regions, and scales, Hawkins et al. (2007) attempt a more comprehensive test of predictions made by proponents of the metabolic theory of ecology about the relationship of temperature and species richness. The predictions are that temperature is the dominant factor controlling species richness patterns, and therefore, that observed species richness will scale log-linearly with (rescaled) temperature with a slope of about −0.65 (Allen et al. 2003, Brown et al. 2004). A wide-ranging survey is necessary to assess such general predictions, but also raises a familiar challenge: what is the best way to pool information across diverse data sets and to reach an overall conclusion with the appropriate degree of confidence? Here, I argue that a hierarchical statistical framework provides a flexible, robust way of dealing with precisely these issues. By reanalyzing a subset of the data in such a framework, I confirm the main findings of the paper: that terrestrial richness patterns do not generally conform to the MTE's predictions. The reanalysis also reveals a surprising association between the slope inferred for a data set and the data set's latitudinal extent, showing that temperature cannot be the sole important driver of terrestrial species richness patterns. MTE's predictions fail in this case probably for two reasons. (1) At high temperatures, other resources inversely correlated with temperature (mostly water) control species richness, often producing a decline in species richness at the highest temperatures and, hence, a shallower slope or nonlinearity. (2) Dispersal of organisms from their place of speciation tends dilute the richness–temperature relationship throughout its range, producing shallower slopes.

To use a collection of data such as the impressive one presented here, it is necessary to synthesize the multiple data sets and arrive at an overall conclusion. The authors score individual data sets as "consistent," "possibly consistent," or "inconsistent" with MTE, and then perform a meta-analysis. This approach provides a summary, but the scoring system, although logical, does not take into account differences among the data sets in sample size and informativeness to provide an integrated measure of uncertainty. As a complement to the authors' approach, I reanalyzed a subset of the data, using hierarchical Bayesian regression models, with OpenBUGS 2.0 (Thomas et al. 2006). These models are included as a Supplement. The models fit slopes and intercepts to individual data sets, while allowing these individual slopes to inform an overall common slope and intercept. The models thus provide inference simultaneously about a "consensus" slope to which slopes for the individual data sets are related, and the degree to which individual slopes depart from that consensus slope (Gelman et al. 1995); see Model 1 of the Supplement. The results also quantify the uncertainty around the fitted slopes, so that we can assess the strength of agreement or disagreement of the data with the slope of about −0.65 predicted by MTE.

I limited the reanalysis to the 23 data sets that the authors identify as linear. It would be difficult to interpret slopes fitted through the clearly nonlinear data sets, and without a geographical basis for splitting them into subgroups and lacking the authors' intimate familiarity with the data, interpretation of fits to partial data sets is also not straightforward. A disadvantage of using the subset is that it removes some of the data sets with broadest geographical coverage, and thus weakens the conclusions as to richness patterns at the global scale. On the other hand, these large data sets are obviously nonlinear and are therefore not consistent with MTE, so arguably they ought not to be used to test the more precise prediction of the value of the linear slope. The reanalysis does not use RMA (reduced major axis) regression, so it might be criticized for sensitivity to error in the explanatory variable (temperature). As an alternative to the all-or-nothing choice between OLS (ordinary least squares) and RMA regression, it is straightforward in the Bayesian framework to add a

Manuscript received 17 November 2006; revised 21 December 2006; accepted 22 December 2006; final version received 15 January 2007. Corresponding Editor: A. M. Ellison.

[1] E-mail: andrew.latimer@gmail.com

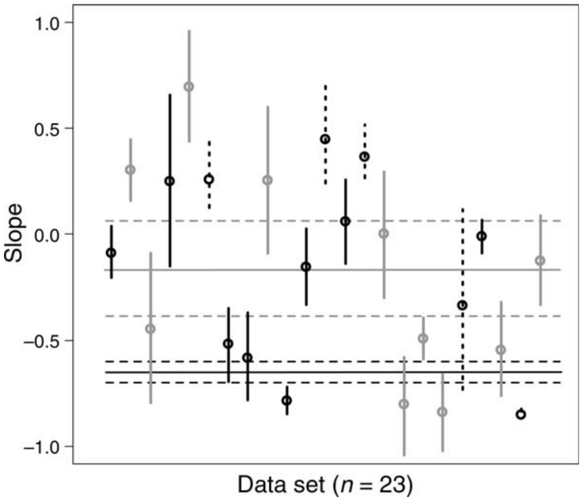

FIG. 1. Means and credible intervals for the slope parameters of 23 linear data sets. Each point represents the posterior mean of the slope for one data set in the hierarchical analysis, and the vertical lines span the 95% credible interval for the slope. The different line types indicate broad taxonomic groupings: black solid, vertebrates; gray solid, insects; and black dashed, plants. The black horizontal lines indicate the slope predictions of the metabolic theory of ecology (MTE): the solid line is at −0.65, and the dashed black lines are at −0.6 and −0.7. The gray horizontal lines display the hierarchical model result for the common slope: the solid gray line is the posterior mean, and the dashed gray lines are the 0.025 and 0.975 quantiles.

submodel for errors in variables. For example, temperature observations can be modeled as realizations of a process with normal errors, using a fitted variance parameter for the errors (see Model 2, Supplement). This modification did not substantially affect any of the slope estimates, so the temperature–richness slope results presented here are from the simpler model (Model 1, Supplement).

My results are generally consistent with those of Hawkins et al. (2007). First, the overall "consensus" slope for the linear data sets is −0.166, and the 95% credible interval contains 0, extending from −0.384 to 0.062. Second, 16 of the 23 linear data sets have slope estimates with credible intervals excluding the range of slopes predicted by MTE (−0.7 to −0.6), and only five of the 23 slopes (those for Australian tiger beetles, Chinese amphibians and reptiles, Mexican hawk moths, and South African plants) have credible intervals including −0.65, despite the rather wide intervals (Fig. 1). This confirms the authors' conclusions that richness slopes for terrestrial organisms vary widely, are often inconsistent with MTE predictions, and are overall shallower than predicted by MTE.

What accounts for the variation in slopes and the frequent nonlinearity of the relationship between predicted metabolic rate and species richness? There is no obvious taxonomic pattern: as Fig. 1 shows, slopes for each broad taxonomic group (vertebrates, insects, and plants) can span the range from less than −0.7 to positive. Surprisingly, there is a strong negative relationship between the latitudinal extent of the data sets and their slopes. As Fig. 2 shows, the data sets that span the largest latitudinal extents have the most negative slopes. One of the advantages of the hierarchical modeling framework is that covariates can be included into the structure to assess whether they contribute significantly to explaining the observed pattern (see Model 3, Supplement). When latitudinal range was included as a covariate in the hierarchical model, its coefficient had a significantly negative value (mean = −0.333, 95% credible interval from −0.526 to −0.132). Strikingly, the temperature range covered by data sets did not significantly affect slope (the coefficient for temperature range had a mean of −0.042 and credible interval from −0.436 to 0.396), and latitudinal range remained significant when included in the model with temperature range. Failure of many data sets to conform to the MTE prediction cannot, then, be dismissed as the result of sampling temperatures too narrowly.

There is a second noteworthy latitudinal range effect in the data. Compared across all 46 data sets, not just the linear subset discussed above, the data sets that the authors found to be nonlinear had significantly larger mean latitudinal extents than the linear data sets (two-tailed *t* test, $t = -4.99$, $P < 0.001$). Nonlinearity in the 46 data sets was also strongly associated with temperature range (two-tailed *t* test, $t = -5.38$, $P < 0.001$).

The MTE cannot explain why slope should depend on latitudinal extent, unless latitudinal extent is strongly correlated with temperature range, because there are no geographical parameters in the model. In data sets covering only a limited range of temperatures, noise might be expected to obscure any richness signal (Brown

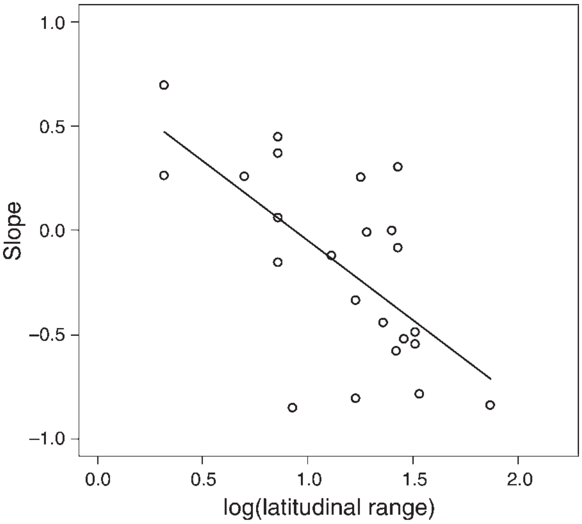

FIG. 2. Scatter plot of inferred slopes against log-transformed latitudinal extent (range in degrees) for the 23 linear data sets. The trend line is the least-squares linear fit through the points.

et al. 2004). However, in the data analyzed here, data sets that span a wide temperature range are more likely to show a nonlinear richness response, and temperature range only marginally affects slope. Rather, it is the strictly geographical factor, latitudinal extent, that determines the nature of the relationship of richness to temperature. This result is inconsistent with MTE, because this geographic factor appears to be swamping local temperature in determining richness patterns for terrestrial species.

The more important question is why. There are two kinds of explanations for inconsistencies between MTE and the data: (1) factors that affect the relationship between temperature and speciation rate, such as co-limiting resources; and (2) factors that affect the relationship between speciation rate and observed richness patterns, such as dispersal.

Temperature and speciation rate: water limitation

One of the MTE's equations relating speciation rate to temperature is

$$K = [R]M^{-3/4}e^{-E/kT}$$

which specifies how carrying capacity K, and thus the density of organisms, is related to body mass M, temperature T, activation energy E, and resource availability $[R]$ (Brown et al. 2004). Underlying the prediction that speciation rate is strongly linked to temperature is the assumption that $[R]$ is not limited by some resource that is independent of, or negatively correlated with, T (Sterner 2004). If this assumption is violated, such that another resource such as water limits K and is inversely correlated with T, the temperature–richness slope will be shallower, or even positive, as observed here.

In terrestrial systems, water may limit productivity and often varies inversely with temperature on local to regional scales due to orographic and coastal rainfall effects. The authors point out that water limitation may play a key role in producing shallower and even positive slopes; data sets from summer-dry regions in which water generally limits productivity (e.g., Iberia, Colorado/Nevada, Australia) tend to have positive slopes (see Hawkins et al. 2007: Table 1). This hypothesis is also consistent with preliminary results on species richness along elevational gradients, which suggest there is frequently a mid-elevation peak in diversity, particularly in drier areas (Rahbek 2005, Kluge et al. 2006).

The hierarchical model provides a framework for a preliminary test of this hypothesis. I obtained data on annual precipitation and precipitation in the driest quarter from WorldClim (Hijmans et al. 2005) for the latitude/longitude locations associated with the 23 linear data sets, to produce a mean value for each data set. When warm-season precipitation is included as a covariate, it has a negative, although marginally nonsignificant, relationship with slope (mean = -0.23, 95% credible interval from -0.46 to 0.024), confirming that there is a trend for data sets from summer-dry regions to have less negative slopes.

Speciation rate and richness: dispersal effects

The fate of species after speciation is not integral to MTE itself, and perhaps for that reason it is only briefly mentioned in the papers proposing the temperature–richness link (Allen et al. 2003, Brown et al. 2004). However, it is likely that dynamics affecting distributions of species after they arise will strongly interfere with richness patterns. At the most basic level, dispersal of species away from their site of origin will tend to reduce the richness–temperature slope (by adding species to the cooler regions). Such "leakage" might be greatest where domains are small and a single dispersal event can move an individual across the domain. Non-negligible levels of dispersal will raise species richness levels in cooler regions above what they would be if all species originated locally, producing a shallower richness–temperature slope. This observation also produces a testable hypothesis: for groups of organisms that disperse well, the relationship between temperature and richness should be weaker than for poorly dispersed groups.

By contrast, for the MTE to predict richness slopes on elevational gradients correctly, species must be environmentally limited but disperse well so that the species sample on any particular mountain is a good sample of the regional species pool. Note that if species disperse well enough to reach every elevational band where they could occur, they are also likely to disperse well enough to get to regions beyond the temperatures where speciation is occurring, so that there is likely to be a tension between processes favoring the MTE's predictions on elevational gradients and those favoring its predictions on latitudinal gradients.

CONCLUSION

For these reasons, it is not surprising that terrestrial richness patterns do not conform to MTE predictions. These findings do not discredit MTE as a conceptual approach, of course, or affect the theory's predictions in other areas of ecology. Even regarding species diversity, the idea of linking total metabolic activity and generation time to speciation rates may still prove useful. For example, this relationship might underlie a general relationship between richness and productivity. But it appears that an adequate model of terrestrial species richness, whether based on metabolism or not, will have to take account of more than one driving factor. The ability of such a model to predict richness patterns will depend to some degree on extra-metabolic factors such as spatial scale and dispersal ability. The next step will be to use explicit comparisons among different kinds of organisms, biomes, and scales to assess the importance of such mechanisms in affecting species richness, preferably extending to other kinds of environments, such as marine systems, and to smaller organisms such as plankton and bacteria.

ACKNOWLEDGMENTS

I thank the authors for generously making their data available for reanalysis, Rob Dunn for helpful insights, and two anonymous reviewers for their very useful comments. NSF grant DEB 0516320 provided funding.

LITERATURE CITED

Allen, A. P., J. F. Gillooly, and J. H. Brown. 2003. Global biodiversity, biochemical kinetics, and the energetic-equivalence rule. Science 297:1545–1548.

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

Gelman, A., J. B. Carlin, and D. B. Rubin. 1995. Bayesian data analysis. Chapman and Hall, New York, New York, USA.

Hawkins, B. A., et al. 2007. A global evaluation of metabolic theory as an explanation for terrestrial species richness gradients. Ecology 88:1877–1888.

Hijmans, R. J., S. E. Cameron, J. L. Parra, P. G. Jones, and A. Jarvis. 2005. Very high resolution interpolated climate surfaces for global land areas. International Journal of Climatology 25:1965–1978.

Kluge, J., M. Kessler, and R. R. Dunn. 2006. What drives elevational patterns of diversity? A test of geometric constraints, climate and species pool effects for pteridophytes on an elevational gradient in Costa Rica. Global Ecology and Biogeography 15:358–371.

Rahbek, C. 2005. The role of spatial scale and the perception of large-scale species-richness patterns. Ecology Letters 8:224–239.

Sterner, R. W. 2004. A one-resource "stoichiometry"? Ecology 85:1813–1816.

Thomas, A., B. O'Hara, U. Ligges, and S. Sturtz. 2006. Making BUGS Open. R News 6:12–17.

SUPPLEMENT

OpenBUGS code for statistical models (*Ecological Archives* E088-113-S1).

Ecology, 88(8), 2007, pp. 1898–1902

METABOLIC THEORY AND DIVERSITY GRADIENTS: WHERE DO WE GO FROM HERE?

BRADFORD A. HAWKINS,[1,9] JOSÉ ALEXANDRE FELIZOLA DINIZ-FILHO,[2] LUIS MAURICIO BINI,[2] MIGUEL B. ARAÚJO,[3]
RICHARD FIELD,[4] JOAQUÍN HORTAL,[3] JEREMY T. KERR,[5] CARSTEN RAHBEK,[6] MIGUEL Á. RODRÍGUEZ,[7] AND
NATHAN J. SANDERS[8]

[1]*Department of Ecology and Evolutionary Biology, University of California, Irvine, California 92697 USA*
[2]*Departamento de Biologia Geral, ICB, Universidade Federal de Goiás, CP 131, 74.001-970, Goiânia, GO, Brazil*
[3]*Departamento de Biodiversidad y Biología Evolutiva, Museo Nacional de Ciencias Naturales (CSIC), 28006 Madrid, Spain*
[4]*School of Geography, University of Nottingham NG7 2RD UK*
[5]*Department of Biology, University of Ottawa, Ottawa, Ontario K1N 6N5 Canada*
[6]*Center for Macroecology, Institute of Biology, University of Copenhagen, DK-2100 Copenhagen, Denmark*
[7]*Departamento de Ecología, Universidad de Alcalá, 28871 Alcalá de Henares, Madrid, Spain*
[8]*Department of Ecology and Evolutionary Biology, University of Tennessee, Knoxville, Tennessee 37996 USA*

INTRODUCTION

Evaluating the Metabolic Theory of Ecology (MTE) of Brown et al. (2004) with respect to broadscale diversity gradients (see Allen et al. 2002, 2006) was the motivation of Hawkins et al. (2007). We used 46 data sets to test predictions for the slope(s) describing the relationship between species richness and temperature. The predicted slopes were found in very few data sets, leading us to question MTE as a general framework for understanding terrestrial diversity gradients. Latimer (2007) reanalyzes some of our data sets using a Bayesian approach and supports our conclusions, whereas Gillooly and Allen (2007) [hereafter G&A] disagree with our approach and raise a number of epistemological issues regarding our evaluation of MTE. Here, we address these issues, focusing on the structure of theories and how a change in epistemological framework undermines the relative strengths of MTE.

THEORIES, HYPOTHESES, AND MODELS

We view MTE as a general *theory*, defined as "a logical construction comprising propositions, some of which contain established information (axioms) while others define questions (postulates). The working part of

---

Manuscript received 23 December 2006; accepted 4 January 2007. Corresponding Editor: A. M. Ellison.

[9] E-mail: bhawkins@uvi.edu

a theory provides the information and logical basis for making generalizations" (Ford 2002:43).

From a body of knowledge encompassed by a theory, *postulates* are derived that must be investigated to support the theory's generality (Ford 2002). Allen et al. (2002:1545) established one postulate, stating unambiguously that their extension of MTE "quantitatively predicts how species diversity increases with environmental temperature." Such clarity is rare among theories purporting to explain broadscale diversity gradients (but see Field et al. [2005]). Allen et al. (2002) and subsequently Brown et al. (2004) also presented their *hypothesis* for diversity gradients as a formal *model*, proposing that the relationship between ln-transformed richness and $1/kT$ (where k is Boltzman's constant and T is temperature in kelvins) has a negative relationship with a slope between −0.6 and −0.7 (in the 2004 version of the model). They also made numerous *data statements*, which define the scientific procedure for investigating a postulate by specifying the measurements to be taken, the data requirements, and the statistical tests to be applied (Ford 2002).

Allen et al. (2002) tested their model using seven data sets comprising both altitudinal and latitudinal gradients. Hawkins et al. (2007) simply expanded this test to a large number of broadscale data sets selected solely on the basis of data availability. The results were inconsistent with MTE predictions in most cases (see also Algar et al. 2007). In response, G&A claim that we oversimplified the theory and used the wrong methodology. However, we used the methods developed by Allen et al. (2002), and the model that we tested was exactly as described by Allen et al. (2002) and Brown et al. (2004).

G&A also argue that we misunderstand Allen et al. (2002), because we ignored later developments in MTE that provide an evolutionary and mechanistic basis for the theory (e.g., Allen et al. 2006). We strongly support evolutionary approaches to understanding diversity gradients (see, e.g., Hawkins et al. 2005, 2006, *in press*), but the newer models must be the subject of future tests. For now, we cannot find where in Allen et al. (2006) or G&A these new developments are said to invalidate Allen et al. (2002) and Brown et al. (2004), so we have to assume that slopes between −0.6 and −0.7 still constitute a valid prediction of their models. Adding a mechanism to a model based on the same theory should not change the basic patterns predicted by the model, unless one or the other is intrinsically wrong or incorrectly developed. As the more recent papers provide no new prediction for the relationship between richness and temperature, it is unclear how the new developments invalidate the conclusions of Hawkins et al. (2007). Alternatively, if the new work shows that the original prediction was not correct, then we agree that the MTE model presented in Allen et al. (2002) and Brown et al. (2004) is not an unequivocal explanation for diversity gradients. This leads us to the next issue: what assumptions must be met and what data statements are necessary to test a theory?

Assumptions and Data Statements

To test a theory we first need to know when and where it applies. Clearly, MTE applies to ectotherms, but G&A say that we should exclude many groups of ectotherm organisms that are able to "maintain relatively constant body temperatures in different thermal environments," and thus the model only applies to "true ectotherms." Also, it is "not expected for groups that are narrowly defined" (G&A). Further, Allen et al. (2002:1547) say, "...we do not mean to imply that temperature is the only variable that affects biodiversity," which G&A reiterate. They recognize that other factors are important (see also Whittaker et al. 2001, Willig et al. 2003), and their stated purpose was to "only predict the slope of the diversity–temperature plots" (Allen et al. 2002:1547). This was also the purpose of Hawkins et al. (2007). Additional restrictive conditions with respect to MTE's applicability are also advanced by G&A: we should avoid areas with extreme water deficits and regions without a broad range of temperatures, although Latimer (2007) reports that the latter condition does not explain poor model fits. Taken together, the restrictive conditions lead to a revised claim that MTE explains richness gradients when it is not too hot, too dry, the wrong region, the wrong scale, or the wrong group. At this point, it is legitimate to question the scope and generality of the theory.

If a model is built on unrealistic assumptions, empirical data should rarely agree with it. The model of Allen et al. (2002:1546), stating that "the natural logarithm of species richness should be a linear function of $1000/T$" (or $1/kT$ in Brown [2004]), is based on several key assumptions (e.g., communities follow the energetic equivalence rule, and abundance and average body size are spatially invariant). Testing these assumptions thus requires detailed data on variation in body size and abundance at broad spatial scales. It is also difficult to know whether the assumptions are realistic, or how violating them affects the model's predictions (see Currie et al. 2004). G&A question our analysis because the data were not selected carefully to meet all of the assumptions, but it is clear that neither Allen et al. (2002) nor any of the subsequent papers were able to check the assumptions for the data that they used. Our data are at least equivalent to the broadscale data that they and others have used to support MTE. Therefore, if our data are questionable then all published analyses cited by G&A using broadscale data are equally questionable. Proponents should not dismiss nonconfirmatory results based on data quality, unless they subject results claimed to support their model to an equally rigorous evaluation of the data and consideration of underlying assumptions.

G&A's criticisms of our use of some data sets highlight that proponents must be much more explicit about data statements than they have been. We welcome the clarifications that they provide, but additional data statements are still needed. How do ecologists obtain the "correct" data? How should we test MTE predictions in a given situation? It is obvious that data should lie in the model's domain, but these must be clearly defined: which taxonomic groups are appropriate; in what environmental conditions does it apply (e.g., what temperature range and water deficit); which measure of temperature should be used? These issues are critical if they want to generate a formal, testable theory for diversity gradients.

Another key issue regarding data statements concerns statistical methods. For example, should we use model I or model II regression? Proponents' claims are inconsistent on this: compare Allen et al. (2002) and Brown et al. (2004) and note that G&A introduce yet another method. Should we use spatially explicit regression models rather than nonspatial methods, or do these only increase uncertainty when correcting Type I errors due to spatial autocorrelation? Further, because multiple factors interact to affect biodiversity, should we generate models with many variables and use partial regression coefficients for temperature? If so, what variables must be included? Shifting to a multiple regression approach will also mean that multicollinearity will be a potentially serious problem (Graham 2003). Finally, and most importantly, the potential overlap of predictions of MTE and those of alternative models must be considered. This leads to our final point about confronting models with data.

Hypothesis Testing and Model Selection

We agree that MTE initially had an advantage over theories based on purely correlative methods. The attractive feature of the model of Allen et al. (2002) was that it provides a theoretical prediction that can be compared with observed slopes. Testing such predictions is usually done in a Fisherian-Popperian framework. However, G&A argue that this results in "unreasonably casting aside this young theory." Although this epistemological framework may indeed be questioned and alternative frameworks do exist (see Hilborne and Mangel 1997), it is widely accepted that the Fisherian-Popperian framework permits "strong" tests in ecology, as opposed to weak tests based on inductive curve fitting (see McGill 2003).

Hawkins et al. (2007) compared observed and predicted slopes using 46 data sets, further dividing nonlinear data into pieces to increase the chances of finding supportive slopes in regions where energy is expected to influence diversity strongly (Hawkins et al. 2003, Whittaker et al. 2007). Although many 95% CI intervals encompassed the predicted slopes, they also encompassed zero, giving the null hypothesis of no relationship between richness and temperature equal standing from a hypothesis-testing perspective. Further, the distribution of slopes was extremely broad and centered nowhere near −0.65. Ultimately, using OLS regression, only one of the 46 data sets was consistent with the coupled predictions of Allen et al. (2002) and Brown et al. (2004) that the relationship between rescaled temperature and ln-transformed richness is both linear and has a slope near −0.65 (none were consistent using RMA regression). G&A accuse us of being too Popperian, but an acceptance rate of 0–2% offers minimal support for a hypothesis under any framework and casts serious doubt about the validity of the postulate. To sidestep this, G&A recommend a shift from a falsificatory to a confirmatory testing procedure. This is in part what Latimer (2007) did using a Bayesian approach, by finding a "consensus" slope for 23 of our data sets instead of testing individual slopes against the predicted value of −0.65. It is important to note that Hawkins et al. (2007) also used a similar approach by performing a meta-analysis for the same purpose, with results that were largely confirmed by Latimer's (2007) reanalysis. Even so, switching tests of MTE from a falsificatory to a confirmatory procedure also creates new problems, to which we now turn.

If predictions of MTE become vague and not subject to falsification, how does MTE differ from other theories (see Lavers and Field 2006)? G&A optimistically interpret our results as promising, despite the extreme range of slopes found. They note that, after controlling for the effects of other variables, one data set shows an "exponential increase of richness with temperature," arguing that this is consistent with the model of Allen et al. (2002). But it may also be consistent with most theories for geographical diversity gradients, highlighting the limitation of the confirmatory approach when multiple models make qualitatively similar predictions. We also consider a defense of MTE based on the "youth" of the theory to be an a posteriori attempt to salvage it after its central predictions fail. Proponents should abandon the "baby in the bathwater" argument in either a falsificatory or a confirmatory epistemological context.

Using a confirmatory approach, G&A nonrandomly select three of our 46 data sets for reanalysis, but instead of fitting the best model under least squares, they force a slope of −0.65 and interpret the explanatory power of their model based on coefficients of determination. Notably, one of the groups that they selected (tiger beetles) is inconsistent with two of their restrictive conditions, being a narrowly defined taxonomic group and comprising species that thermoregulate (Pearson and Vogler 2001, Dajoz 2002). They also select amphibians, but many of these also thermoregulate (Hutchinson and Dupré 1992). This illustrates the difficulty in understanding when the theory applies. Irrespectively, we repeated their approach for all 46 data sets, ignoring any nonlinearity following G&A but violating the postulate of linearity by Allen et al. (2002). The coefficients of determination of these tests were very low, with 27 being zero, and eight others being less than 0.30 (Table 1). Across all data sets, the r^2 values were substantially lower than the r^2 values from OLS fits (paired t test = −5.39; $P < 0.001$), despite low overall fits of temperature using either method (average $r^2_{\text{G\&A}} =$

TABLE 1. Coefficients of determination (r^2) for linear regressions of ln-transformed richness against rescaled temperature using ordinary least squares [OLS] vs. the "forced slope" method of Gillooly and Allen (2007) [G&A].

Group	Region	OLS	G&A
Amphibians	Afrotropics	0.216	0
Amphibians	Australia	0.003	0
Amphibians	Brazil	0.590	0
Amphibians	China	0.404	0.384
Amphibians	Europe	0.502	0.499
Amphibians	Iberia	0.010	0
Amphibians	North America	0.767	0.739
Angiosperms	China	0.353	0.059
Ants	Colorado/Nevada	0.053	0
Ants	New World	0.582	0.545
Blister beetles	North America	0.347	0.312
Bumble bees	global	0.165	0
Butterflies	Australia	0.030	0
Butterflies	California	0.306	0
Butterflies	western Palearctic	0.136	0
Butterflies (summer)	North America	0.606	0.261
Butterflies (winter)	North America	0.499	0
Dung beetles	Iberia/France	0.008	0
Dung beetles	western Palearctic	0.111	0
Eupelmid wasps	Palearctic	0.084	0
Grasshoppers	North America	0.284	0.016
Hawk moths	Mexico	0.221	0.209
Hawk moths	southeastern Asia	0.025	0
Orthoptera	Catalonia	0.101	0
Plants	California	0.192	0
Plants	Catalonia	0.044	0
Plants (exotic)	Great Britain	0.656	0.190
Plants (native)	Great Britain	0.519	0.490
Pteridophytes	Europe	0.208	0
Pteridophytes	Iberia	0.055	0
Reptiles	Brazil	0.014	0
Reptiles	China	0.383	0.381
Reptiles	Europe	0.607	0.588
Reptiles	Iberia	0.002	0
Reptiles	North America	0.838	0.620
Reptiles	southern Africa	0	0
Seed plants	Iberia	0.082	0
Snakes	Afrotropics	0.278	0
Tiger beetles	Australia	0.113	0.099
Tiger beetles	India	0	0
Tiger beetles	North America	0.560	0.544
Tiger beetles	northwestern South America	0.156	0.153
Trees	Europe	0.458	0.372
Trees	North America	0.588	0.584
Woody plants	Kenya	0.338	0
Woody plants	southern Africa	0.019	0.012

Note: Regressions were done across all values within each of the 46 data sets ignoring any nonlinearity in the data.

0.153; average $r^2_{\mathrm{OLS}} = 0.272$). Although we currently do not have other environmental predictors for all data sets, previous meta-analyses (Hawkins et al. 2003) indicate that r^2 values of other variables (derived from theories related to water–energy balance; e.g., O'Brien [2006]) have much greater statistical explanatory power. Moreover, recent modeling of geographic range overlap explicitly based on MTE generated results with lower explanatory power than those generated using alternative models (Rahbek et al. 2007).

If the confirmatory approach is to be used for testing MTE, and any positive relationship between temperature and diversity is "promising," evaluations will become mainly correlative, as with many competing theories. Therefore, model developers must clearly describe the unique predictions made by their model (Shipley 2000, Currie et al. 2004). This is essential for understanding diversity gradients, because the spatial structure of climatic variation on Earth causes nearly all theories developed to explain broadscale richness gradients to predict a positive correlation between richness and temperature, even when no causal link between them exists, such as in the "pure tropical conservatism" model (Wiens and Donoghue 2004).

CONCLUDING REMARKS

MTE can be viewed as the core of a research program. The hypothesis of Allen et al. (2002), together with the model(s) developed to test it, is one facet of the program. Their model(s) can be tested and rejected, but this does not necessarily challenge the core. As pointed out by Hawkins et al. (2007), our evaluation was restricted to the predictions of Allen et al. (2002) and Brown et al. (2004) for richness gradients and cannot be generalized to MTE as a whole (also see Latimer 2007). Even so, we contend that the tests by Hawkins et al. (2007) are as valid as proponents' tests and provide strong evidence against the model as a *general* explanation. Of course, it is difficult to know whether the failure of the model's predictions occurs at the postulate, hypothesis, or theory level. Incorporating additional variables (including spatial variation in average body size and abundance, as well including potential deviations from the energetic equivalence rule) might generate improved models that better fit the empirical data. Perhaps this could support the claim that MTE explains richness gradients, at least in part (see also Latimer 2007). But arguing that it might and showing to what extent it does are very different propositions.

LITERATURE CITED

Algar, A. C., J. T. Kerr, and D. J. Currie. 2007. A test of metabolic theory as the mechanism underlying broad-scale species-richness gradients. Global Ecology and Biogeography 16:170–178 [doi: 10.1111/j.1466-8238.2006.oo275x].

Allen, A. P., J. F. Gillooly, and J. H. Brown. 2002. Global biodiversity, biochemical kinetics, and the energetic-equivalence rule. Science 297:1545–1548.

Allen, A. P., J. F. Gillooly, V. M. Savage, and J. H. Brown. 2006. Kinetic effects of temperature on rates of genetic divergence and speciation. Proceedings of the National Academy of Sciences (USA) 103:9130–9135.

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

Currie, D. J., G. G. Mittelbach, H. V. Cornell, R. Field, J.-F. Guégan, B. A. Hawkins, D. M. Kaufman, J. T. Kerr, T. Oberdorff, E. O'Brien, and J. R. G. Turner. 2004. Predictions and tests of climate-based hypotheses of broad-scale variation in taxonomic richness. Ecology Letters 7:1121–1134.

Dajoz, R. 2002. Les coléoptères carabidés et ténébrionidés: écologie et biologie. Tec & Doc, Paris, France.

Field, R., E. M. O'Brien, and R. J. Whittaker. 2005. Global models for predicting woody plant richness from climate: development and evaluation. Ecology 86:2263–2277.

Ford, E. D. 2002. Scientific method for ecological research. Cambridge University Press, Cambridge, UK.

Gillooly, J. F., and A. P. Allen. 2007. Linking global patterns in biodiversity to evolutionary dynamics using metabolic theory. Ecology 88:1890–1894.

Graham, M. H. 2003. Confronting multicollinearity in ecological multiple regression. Ecology 84:2809–2815.

Hawkins, B. A., et al. 2007. A global evaluation of metabolic theory as an explanation for terrestrial species richness gradients. Ecology 88:1877–1888.

Hawkins, B. A., J. A. F. Diniz-Filho, C. A. Jaramillo, and S. A. Soeller. 2006. Post-Eocene climate change, niche conservatism, and the latitudinal diversity gradient of New World birds. Journal of Biogeography 33:770–780.

Hawkins, B. A., J. A. F. Diniz-Filho, C. A. Jaramillo, and S. A. Soeller. *In press*. Climate, niche conservatism, and the global bird diversity gradient. American Naturalist.

Hawkins, B. A., J. A. F. Diniz-Filho, and S. A. Soeller. 2005. Water links the historical and contemporary components of the Australian bird diversity gradient. Journal of Biogeography 32:1035–1042.

Hawkins, B. A., R. Field, H. V. Cornell, D. J. Currie, J. F. Guégan, D. M. Kaufman, J. T. Kerr, G. G. Mittelbach, T. Oberdorff, E. M. O'Brien, E. E. Porter, and J. R. G. Turner. 2003. Energy, water, and broad-scale geographic patterns of species richness. Ecology 84:3105–3117.

Hilborne, R., and M. Mangel. 1997. The ecological detective: confronting models with data. Princeton University Press, Princeton, New Jersey, USA.

Hutchinson, V. H., and R. K. Dupré. 1992. Thermoregulation. Pages 206–249 *in* M. E. Feder and W. W. Burggren, editors. Environmental physiology of the amphibians. University of Chicago Press, Chicago, Illinois, USA.

Latimer, A. M. 2007. Geography and resource limitation complicate metabolism-based predictions of species richness. Ecology 88:1895–1898.

Lavers, C., and R. Field. 2006. A resource-based conceptual model of plant diversity that reassesses causality in the productivity–diversity relationship. Global Ecology and Biogeography 15:213–224.

McGill, B. 2003. Strong and weak tests of macroecological theory. Oikos 102:679–685.

O'Brien, E. M. 2006. Biological relativity to water–energy dynamics. Journal of Biogeography 33:1868–1888.

Pearson, D. L., and A. P. Vogler. 2001. Tiger beetles: the evolution, ecology, and diversity of the cicindelids. Cornell University Press, Ithaca, New York, USA.

Rahbek, C., N. J. Gotelli, R. K. Colwell, G. L. Entsminger, T. F. L. V. B. Rangel, and G. R. Graves. 2007. Predicting continental-scale patterns of bird species richness with spatially explicit models. Proceedings of the Royal Society B 274:165–174 [doi: 10.1098/rspb.2006.3700].

Shipley, B. 2000. Cause and correlation in biology. Cambridge University Press, Cambridge, UK.

Whittaker, R. J., D. Nogués-Bravo, and M. B. Araújo. 2007. Geographic gradients of species richness: a test of the water–energy conjecture of Hawkins et al. (2003) using European data for five taxa. Global Ecology and Biogeography 16:76–89.

Whittaker, R. J., K. J. Willis, and R. Field. 2001. Scale and species richness: towards a general, hierarchical theory of species diversity. Journal of Biogeography 28:453–470.

Wiens, J. J., and M. J. Donoghue. 2004. Historical biogeography, ecology and species richness. Trends in Ecology and Evolution 19:639–644.

Willig, M. R., D. M. Kaufman, and R. D. Stevens. 2003. Latitudinal gradients of biodiversity: pattern, process, scale and synthesis. Annual Review of Ecology, Evolution and Systematics 34:273–309.

Ecology, 88(8), 2007, pp. 1895–1898

GEOGRAPHY AND RESOURCE LIMITATION COMPLICATE METABOLISM-BASED PREDICTIONS OF SPECIES RICHNESS

Andrew M. Latimer[1]

Department of Ecology and Evolutionary Biology, University of Connecticut, 75 North Eagleville Road, Storrs, Connecticut 06269 USA

Metabolic Theory of Ecology: A Reanalysis of Data

Assessments of general theoretical frameworks like neutral ecological theory and the metabolic theory of ecology (MTE) often use only one or two data sets, and thus do not provide much evidence about the particular prediction tested, much less the validity or usefulness of the general theory. By compiling a diverse group of data sets from a range of terrestrial taxa, regions, and scales, Hawkins et al. (2007) attempt a more comprehensive test of predictions made by proponents of the metabolic theory of ecology about the relationship of temperature and species richness. The predictions are that temperature is the dominant factor controlling species richness patterns, and therefore, that observed species richness will scale log-linearly with (rescaled) temperature with a slope of about −0.65 (Allen et al. 2003, Brown et al. 2004). A wide-ranging survey is necessary to assess such general predictions, but also raises a familiar challenge: what is the best way to pool information across diverse data sets and to reach an overall conclusion with the appropriate degree of confidence? Here, I argue that a hierarchical statistical framework provides a flexible, robust way of dealing with precisely these issues. By reanalyzing a subset of the data in such a framework, I confirm the main findings of the paper: that terrestrial richness patterns do not generally conform to the MTE's predictions. The reanalysis also reveals a surprising association between the slope inferred for a data set and the data set's latitudinal extent, showing that temperature cannot be the sole important driver of terrestrial species richness patterns. MTE's predictions fail in this case probably for two reasons. (1) At high temperatures, other resources inversely correlated with temperature (mostly water) control species richness, often producing a decline in species richness at the highest temperatures and, hence, a shallower slope or nonlinearity. (2) Dispersal of organisms from their place of speciation tends dilute the richness–temperature relationship throughout its range, producing shallower slopes.

To use a collection of data such as the impressive one presented here, it is necessary to synthesize the multiple data sets and arrive at an overall conclusion. The authors score individual data sets as "consistent," "possibly consistent," or "inconsistent" with MTE, and then perform a meta-analysis. This approach provides a summary, but the scoring system, although logical, does not take into account differences among the data sets in sample size and informativeness to provide an integrated measure of uncertainty. As a complement to the authors' approach, I reanalyzed a subset of the data, using hierarchical Bayesian regression models, with OpenBUGS 2.0 (Thomas et al. 2006). These models are included as a Supplement. The models fit slopes and intercepts to individual data sets, while allowing these individual slopes to inform an overall common slope and intercept. The models thus provide inference simultaneously about a "consensus" slope to which slopes for the individual data sets are related, and the degree to which individual slopes depart from that consensus slope (Gelman et al. 1995); see Model 1 of the Supplement. The results also quantify the uncertainty around the fitted slopes, so that we can assess the strength of agreement or disagreement of the data with the slope of about −0.65 predicted by MTE.

I limited the reanalysis to the 23 data sets that the authors identify as linear. It would be difficult to interpret slopes fitted through the clearly nonlinear data sets, and without a geographical basis for splitting them into subgroups and lacking the authors' intimate familiarity with the data, interpretation of fits to partial data sets is also not straightforward. A disadvantage of using the subset is that it removes some of the data sets with broadest geographical coverage, and thus weakens the conclusions as to richness patterns at the global scale. On the other hand, these large data sets are obviously nonlinear and are therefore not consistent with MTE, so arguably they ought not to be used to test the more precise prediction of the value of the linear slope. The reanalysis does not use RMA (reduced major axis) regression, so it might be criticized for sensitivity to error in the explanatory variable (temperature). As an alternative to the all-or-nothing choice between OLS (ordinary least squares) and RMA regression, it is straightforward in the Bayesian framework to add a

Manuscript received 17 November 2006; revised 21 December 2006; accepted 22 December 2006; final version received 15 January 2007. Corresponding Editor: A. M. Ellison.

[1] E-mail: andrew.latimer@gmail.com

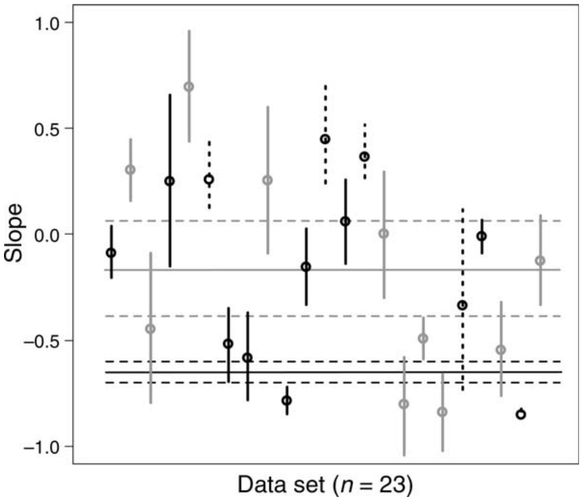

FIG. 1. Means and credible intervals for the slope parameters of 23 linear data sets. Each point represents the posterior mean of the slope for one data set in the hierarchical analysis, and the vertical lines span the 95% credible interval for the slope. The different line types indicate broad taxonomic groupings: black solid, vertebrates; gray solid, insects; and black dashed, plants. The black horizontal lines indicate the slope predictions of the metabolic theory of ecology (MTE): the solid line is at −0.65, and the dashed black lines are at −0.6 and −0.7. The gray horizontal lines display the hierarchical model result for the common slope: the solid gray line is the posterior mean, and the dashed gray lines are the 0.025 and 0.975 quantiles.

submodel for errors in variables. For example, temperature observations can be modeled as realizations of a process with normal errors, using a fitted variance parameter for the errors (see Model 2, Supplement). This modification did not substantially affect any of the slope estimates, so the temperature–richness slope results presented here are from the simpler model (Model 1, Supplement).

My results are generally consistent with those of Hawkins et al. (2007). First, the overall "consensus" slope for the linear data sets is −0.166, and the 95% credible interval contains 0, extending from −0.384 to 0.062. Second, 16 of the 23 linear data sets have slope estimates with credible intervals excluding the range of slopes predicted by MTE (−0.7 to −0.6), and only five of the 23 slopes (those for Australian tiger beetles, Chinese amphibians and reptiles, Mexican hawk moths, and South African plants) have credible intervals including −0.65, despite the rather wide intervals (Fig. 1). This confirms the authors' conclusions that richness slopes for terrestrial organisms vary widely, are often inconsistent with MTE predictions, and are overall shallower than predicted by MTE.

What accounts for the variation in slopes and the frequent nonlinearity of the relationship between predicted metabolic rate and species richness? There is no obvious taxonomic pattern: as Fig. 1 shows, slopes for each broad taxonomic group (vertebrates, insects, and plants) can span the range from less than −0.7 to positive. Surprisingly, there is a strong negative relationship between the latitudinal extent of the data sets and their slopes. As Fig. 2 shows, the data sets that span the largest latitudinal extents have the most negative slopes. One of the advantages of the hierarchical modeling framework is that covariates can be included into the structure to assess whether they contribute significantly to explaining the observed pattern (see Model 3, Supplement). When latitudinal range was included as a covariate in the hierarchical model, its coefficient had a significantly negative value (mean = −0.333, 95% credible interval from −0.526 to −0.132). Strikingly, the temperature range covered by data sets did not significantly affect slope (the coefficient for temperature range had a mean of −0.042 and credible interval from −0.436 to 0.396), and latitudinal range remained significant when included in the model with temperature range. Failure of many data sets to conform to the MTE prediction cannot, then, be dismissed as the result of sampling temperatures too narrowly.

There is a second noteworthy latitudinal range effect in the data. Compared across all 46 data sets, not just the linear subset discussed above, the data sets that the authors found to be nonlinear had significantly larger mean latitudinal extents than the linear data sets (two-tailed t test, $t = -4.99$, $P < 0.001$). Nonlinearity in the 46 data sets was also strongly associated with temperature range (two-tailed t test, $t = -5.38$, $P < 0.001$).

The MTE cannot explain why slope should depend on latitudinal extent, unless latitudinal extent is strongly correlated with temperature range, because there are no geographical parameters in the model. In data sets covering only a limited range of temperatures, noise might be expected to obscure any richness signal (Brown

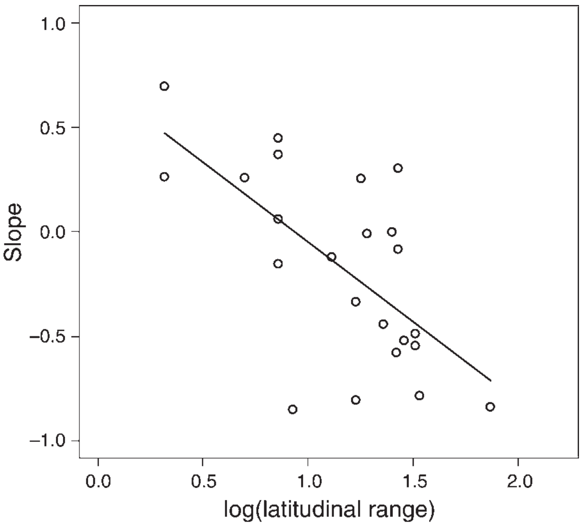

FIG. 2. Scatter plot of inferred slopes against log-transformed latitudinal extent (range in degrees) for the 23 linear data sets. The trend line is the least-squares linear fit through the points.

et al. 2004). However, in the data analyzed here, data sets that span a wide temperature range are more likely to show a nonlinear richness response, and temperature range only marginally affects slope. Rather, it is the strictly geographical factor, latitudinal extent, that determines the nature of the relationship of richness to temperature. This result is inconsistent with MTE, because this geographic factor appears to be swamping local temperature in determining richness patterns for terrestrial species.

The more important question is why. There are two kinds of explanations for inconsistencies between MTE and the data: (1) factors that affect the relationship between temperature and speciation rate, such as co-limiting resources; and (2) factors that affect the relationship between speciation rate and observed richness patterns, such as dispersal.

Temperature and speciation rate: water limitation

One of the MTE's equations relating speciation rate to temperature is

$$K = [R]M^{-3/4}e^{-E/kT}$$

which specifies how carrying capacity K, and thus the density of organisms, is related to body mass M, temperature T, activation energy E, and resource availability $[R]$ (Brown et al. 2004). Underlying the prediction that speciation rate is strongly linked to temperature is the assumption that $[R]$ is not limited by some resource that is independent of, or negatively correlated with, T (Sterner 2004). If this assumption is violated, such that another resource such as water limits K and is inversely correlated with T, the temperature–richness slope will be shallower, or even positive, as observed here.

In terrestrial systems, water may limit productivity and often varies inversely with temperature on local to regional scales due to orographic and coastal rainfall effects. The authors point out that water limitation may play a key role in producing shallower and even positive slopes; data sets from summer-dry regions in which water generally limits productivity (e.g., Iberia, Colorado/Nevada, Australia) tend to have positive slopes (see Hawkins et al. 2007: Table 1). This hypothesis is also consistent with preliminary results on species richness along elevational gradients, which suggest there is frequently a mid-elevation peak in diversity, particularly in drier areas (Rahbek 2005, Kluge et al. 2006).

The hierarchical model provides a framework for a preliminary test of this hypothesis. I obtained data on annual precipitation and precipitation in the driest quarter from WorldClim (Hijmans et al. 2005) for the latitude/longitude locations associated with the 23 linear data sets, to produce a mean value for each data set. When warm-season precipitation is included as a covariate, it has a negative, although marginally nonsignificant, relationship with slope (mean = −0.23, 95% credible interval from −0.46 to 0.024), confirming that there is a trend for data sets from summer-dry regions to have less negative slopes.

Speciation rate and richness: dispersal effects

The fate of species after speciation is not integral to MTE itself, and perhaps for that reason it is only briefly mentioned in the papers proposing the temperature–richness link (Allen et al. 2003, Brown et al. 2004). However, it is likely that dynamics affecting distributions of species after they arise will strongly interfere with richness patterns. At the most basic level, dispersal of species away from their site of origin will tend to reduce the richness–temperature slope (by adding species to the cooler regions). Such "leakage" might be greatest where domains are small and a single dispersal event can move an individual across the domain. Non-negligible levels of dispersal will raise species richness levels in cooler regions above what they would be if all species originated locally, producing a shallower richness–temperature slope. This observation also produces a testable hypothesis: for groups of organisms that disperse well, the relationship between temperature and richness should be weaker than for poorly dispersed groups.

By contrast, for the MTE to predict richness slopes on elevational gradients correctly, species must be environmentally limited but disperse well so that the species sample on any particular mountain is a good sample of the regional species pool. Note that if species disperse well enough to reach every elevational band where they could occur, they are also likely to disperse well enough to get to regions beyond the temperatures where speciation is occurring, so that there is likely to be a tension between processes favoring the MTE's predictions on elevational gradients and those favoring its predictions on latitudinal gradients.

CONCLUSION

For these reasons, it is not surprising that terrestrial richness patterns do not conform to MTE predictions. These findings do not discredit MTE as a conceptual approach, of course, or affect the theory's predictions in other areas of ecology. Even regarding species diversity, the idea of linking total metabolic activity and generation time to speciation rates may still prove useful. For example, this relationship might underlie a general relationship between richness and productivity. But it appears that an adequate model of terrestrial species richness, whether based on metabolism or not, will have to take account of more than one driving factor. The ability of such a model to predict richness patterns will depend to some degree on extra-metabolic factors such as spatial scale and dispersal ability. The next step will be to use explicit comparisons among different kinds of organisms, biomes, and scales to assess the importance of such mechanisms in affecting species richness, preferably extending to other kinds of environments, such as marine systems, and to smaller organisms such as plankton and bacteria.

ACKNOWLEDGMENTS

I thank the authors for generously making their data available for reanalysis, Rob Dunn for helpful insights, and two anonymous reviewers for their very useful comments. NSF grant DEB 0516320 provided funding.

LITERATURE CITED

Allen, A. P., J. F. Gillooly, and J. H. Brown. 2003. Global biodiversity, biochemical kinetics, and the energetic-equivalence rule. Science 297:1545–1548.

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

Gelman, A., J. B. Carlin, and D. B. Rubin. 1995. Bayesian data analysis. Chapman and Hall, New York, New York, USA.

Hawkins, B. A., et al. 2007. A global evaluation of metabolic theory as an explanation for terrestrial species richness gradients. Ecology 88:1877–1888.

Hijmans, R. J., S. E. Cameron, J. L. Parra, P. G. Jones, and A. Jarvis. 2005. Very high resolution interpolated climate surfaces for global land areas. International Journal of Climatology 25:1965–1978.

Kluge, J., M. Kessler, and R. R. Dunn. 2006. What drives elevational patterns of diversity? A test of geometric constraints, climate and species pool effects for pteridophytes on an elevational gradient in Costa Rica. Global Ecology and Biogeography 15:358–371.

Rahbek, C. 2005. The role of spatial scale and the perception of large-scale species-richness patterns. Ecology Letters 8:224–239.

Sterner, R. W. 2004. A one-resource "stoichiometry"? Ecology 85:1813–1816.

Thomas, A., B. O'Hara, U. Ligges, and S. Sturtz. 2006. Making BUGS Open. R News 6:12–17.

SUPPLEMENT

OpenBUGS code for statistical models (*Ecological Archives* E088-113-S1).

Ecology, 88(8), 2007, pp. 1898–1902

METABOLIC THEORY AND DIVERSITY GRADIENTS: WHERE DO WE GO FROM HERE?

BRADFORD A. HAWKINS,[1,9] JOSÉ ALEXANDRE FELIZOLA DINIZ-FILHO,[2] LUIS MAURICIO BINI,[2] MIGUEL B. ARAÚJO,[3] RICHARD FIELD,[4] JOAQUÍN HORTAL,[3] JEREMY T. KERR,[5] CARSTEN RAHBEK,[6] MIGUEL Á. RODRÍGUEZ,[7] AND NATHAN J. SANDERS[8]

[1]*Department of Ecology and Evolutionary Biology, University of California, Irvine, California 92697 USA*
[2]*Departamento de Biologia Geral, ICB, Universidade Federal de Goiás, CP 131, 74.001-970, Goiânia, GO, Brazil*
[3]*Departamento de Biodiversidad y Biología Evolutiva, Museo Nacional de Ciencias Naturales (CSIC), 28006 Madrid, Spain*
[4]*School of Geography, University of Nottingham NG7 2RD UK*
[5]*Department of Biology, University of Ottawa, Ottawa, Ontario K1N 6N5 Canada*
[6]*Center for Macroecology, Institute of Biology, University of Copenhagen, DK-2100 Copenhagen, Denmark*
[7]*Departamento de Ecología, Universidad de Alcalá, 28871 Alcalá de Henares, Madrid, Spain*
[8]*Department of Ecology and Evolutionary Biology, University of Tennessee, Knoxville, Tennessee 37996 USA*

INTRODUCTION

Evaluating the Metabolic Theory of Ecology (MTE) of Brown et al. (2004) with respect to broadscale diversity gradients (see Allen et al. 2002, 2006) was the motivation of Hawkins et al. (2007). We used 46 data sets to test predictions for the slope(s) describing the relationship between species richness and temperature. The predicted slopes were found in very few data sets, leading us to question MTE as a general framework for understanding terrestrial diversity gradients. Latimer (2007) reanalyzes some of our data sets using a Bayesian approach and supports our conclusions, whereas Gillooly and Allen (2007) [hereafter G&A] disagree with our approach and raise a number of epistemological issues regarding our evaluation of MTE. Here, we address these issues, focusing on the structure of theories and how a change in epistemological framework undermines the relative strengths of MTE.

THEORIES, HYPOTHESES, AND MODELS

We view MTE as a general *theory*, defined as "a logical construction comprising propositions, some of which contain established information (axioms) while others define questions (postulates). The working part of

Manuscript received 23 December 2006; accepted 4 January 2007. Corresponding Editor: A. M. Ellison.

[9] E-mail: bhawkins@uvi.edu

a theory provides the information and logical basis for making generalizations" (Ford 2002:43).

From a body of knowledge encompassed by a theory, *postulates* are derived that must be investigated to support the theory's generality (Ford 2002). Allen et al. (2002:1545) established one postulate, stating unambiguously that their extension of MTE "quantitatively predicts how species diversity increases with environmental temperature." Such clarity is rare among theories purporting to explain broadscale diversity gradients (but see Field et al. [2005]). Allen et al. (2002) and subsequently Brown et al. (2004) also presented their *hypothesis* for diversity gradients as a formal *model*, proposing that the relationship between ln-transformed richness and $1/kT$ (where k is Boltzman's constant and T is temperature in kelvins) has a negative relationship with a slope between −0.6 and −0.7 (in the 2004 version of the model). They also made numerous *data statements*, which define the scientific procedure for investigating a postulate by specifying the measurements to be taken, the data requirements, and the statistical tests to be applied (Ford 2002).

Allen et al. (2002) tested their model using seven data sets comprising both altitudinal and latitudinal gradients. Hawkins et al. (2007) simply expanded this test to a large number of broadscale data sets selected solely on the basis of data availability. The results were inconsistent with MTE predictions in most cases (see also Algar et al. 2007). In response, G&A claim that we oversimplified the theory and used the wrong methodology. However, we used the methods developed by Allen et al. (2002), and the model that we tested was exactly as described by Allen et al. (2002) and Brown et al. (2004).

G&A also argue that we misunderstand Allen et al. (2002), because we ignored later developments in MTE that provide an evolutionary and mechanistic basis for the theory (e.g., Allen et al. 2006). We strongly support evolutionary approaches to understanding diversity gradients (see, e.g., Hawkins et al. 2005, 2006, *in press*), but the newer models must be the subject of future tests. For now, we cannot find where in Allen et al. (2006) or G&A these new developments are said to invalidate Allen et al. (2002) and Brown et al. (2004), so we have to assume that slopes between −0.6 and −0.7 still constitute a valid prediction of their models. Adding a mechanism to a model based on the same theory should not change the basic patterns predicted by the model, unless one or the other is intrinsically wrong or incorrectly developed. As the more recent papers provide no new prediction for the relationship between richness and temperature, it is unclear how the new developments invalidate the conclusions of Hawkins et al. (2007). Alternatively, if the new work shows that the original prediction was not correct, then we agree that the MTE model presented in Allen et al. (2002) and Brown et al. (2004) is not an unequivocal explanation for diversity gradients. This leads us to the next issue: what assumptions must be met and what data statements are necessary to test a theory?

Assumptions and Data Statements

To test a theory we first need to know when and where it applies. Clearly, MTE applies to ectotherms, but G&A say that we should exclude many groups of ectotherm organisms that are able to "maintain relatively constant body temperatures in different thermal environments," and thus the model only applies to "true ectotherms." Also, it is "not expected for groups that are narrowly defined" (G&A). Further, Allen et al. (2002:1547) say, "…we do not mean to imply that temperature is the only variable that affects biodiversity," which G&A reiterate. They recognize that other factors are important (see also Whittaker et al. 2001, Willig et al. 2003), and their stated purpose was to "only predict the slope of the diversity–temperature plots" (Allen et al. 2002:1547). This was also the purpose of Hawkins et al. (2007). Additional restrictive conditions with respect to MTE's applicability are also advanced by G&A: we should avoid areas with extreme water deficits and regions without a broad range of temperatures, although Latimer (2007) reports that the latter condition does not explain poor model fits. Taken together, the restrictive conditions lead to a revised claim that MTE explains richness gradients when it is not too hot, too dry, the wrong region, the wrong scale, or the wrong group. At this point, it is legitimate to question the scope and generality of the theory.

If a model is built on unrealistic assumptions, empirical data should rarely agree with it. The model of Allen et al. (2002:1546), stating that "the natural logarithm of species richness should be a linear function of $1000/T$" (or $1/kT$ in Brown [2004]), is based on several key assumptions (e.g., communities follow the energetic equivalence rule, and abundance and average body size are spatially invariant). Testing these assumptions thus requires detailed data on variation in body size and abundance at broad spatial scales. It is also difficult to know whether the assumptions are realistic, or how violating them affects the model's predictions (see Currie et al. 2004). G&A question our analysis because the data were not selected carefully to meet all of the assumptions, but it is clear that neither Allen et al. (2002) nor any of the subsequent papers were able to check the assumptions for the data that they used. Our data are at least equivalent to the broadscale data that they and others have used to support MTE. Therefore, if our data are questionable then all published analyses cited by G&A using broadscale data are equally questionable. Proponents should not dismiss nonconfirmatory results based on data quality, unless they subject results claimed to support their model to an equally rigorous evaluation of the data and consideration of underlying assumptions.

G&A's criticisms of our use of some data sets highlight that proponents must be much more explicit about data statements than they have been. We welcome the clarifications that they provide, but additional data statements are still needed. How do ecologists obtain the "correct" data? How should we test MTE predictions in a given situation? It is obvious that data should lie in the model's domain, but these must be clearly defined: which taxonomic groups are appropriate; in what environmental conditions does it apply (e.g., what temperature range and water deficit); which measure of temperature should be used? These issues are critical if they want to generate a formal, testable theory for diversity gradients.

Another key issue regarding data statements concerns statistical methods. For example, should we use model I or model II regression? Proponents' claims are inconsistent on this: compare Allen et al. (2002) and Brown et al. (2004) and note that G&A introduce yet another method. Should we use spatially explicit regression models rather than nonspatial methods, or do these only increase uncertainty when correcting Type I errors due to spatial autocorrelation? Further, because multiple factors interact to affect biodiversity, should we generate models with many variables and use partial regression coefficients for temperature? If so, what variables must be included? Shifting to a multiple regression approach will also mean that multicollinearity will be a potentially serious problem (Graham 2003). Finally, and most importantly, the potential overlap of predictions of MTE and those of alternative models must be considered. This leads to our final point about confronting models with data.

Hypothesis Testing and Model Selection

We agree that MTE initially had an advantage over theories based on purely correlative methods. The attractive feature of the model of Allen et al. (2002) was that it provides a theoretical prediction that can be compared with observed slopes. Testing such predictions is usually done in a Fisherian-Popperian framework. However, G&A argue that this results in "unreasonably casting aside this young theory." Although this epistemological framework may indeed be questioned and alternative frameworks do exist (see Hilborne and Mangel 1997), it is widely accepted that the Fisherian-Popperian framework permits "strong" tests in ecology, as opposed to weak tests based on inductive curve fitting (see McGill 2003).

Hawkins et al. (2007) compared observed and predicted slopes using 46 data sets, further dividing nonlinear data into pieces to increase the chances of finding supportive slopes in regions where energy is expected to influence diversity strongly (Hawkins et al. 2003, Whittaker et al. 2007). Although many 95% CI intervals encompassed the predicted slopes, they also encompassed zero, giving the null hypothesis of no relationship between richness and temperature equal standing from a hypothesis-testing perspective. Further, the distribution of slopes was extremely broad and centered nowhere near −0.65. Ultimately, using OLS regression, only one of the 46 data sets was consistent with the coupled predictions of Allen et al. (2002) and Brown et al. (2004) that the relationship between rescaled temperature and ln-transformed richness is both linear and has a slope near −0.65 (none were consistent using RMA regression). G&A accuse us of being too Popperian, but an acceptance rate of 0–2% offers minimal support for a hypothesis under any framework and casts serious doubt about the validity of the postulate. To sidestep this, G&A recommend a shift from a falsificatory to a confirmatory testing procedure. This is in part what Latimer (2007) did using a Bayesian approach, by finding a "consensus" slope for 23 of our data sets instead of testing individual slopes against the predicted value of −0.65. It is important to note that Hawkins et al. (2007) also used a similar approach by performing a meta-analysis for the same purpose, with results that were largely confirmed by Latimer's (2007) reanalysis. Even so, switching tests of MTE from a falsificatory to a confirmatory procedure also creates new problems, to which we now turn.

If predictions of MTE become vague and not subject to falsification, how does MTE differ from other theories (see Lavers and Field 2006)? G&A optimistically interpret our results as promising, despite the extreme range of slopes found. They note that, after controlling for the effects of other variables, one data set shows an "exponential increase of richness with temperature," arguing that this is consistent with the model of Allen et al. (2002). But it may also be consistent with most theories for geographical diversity gradients, highlighting the limitation of the confirmatory approach when multiple models make qualitatively similar predictions. We also consider a defense of MTE based on the "youth" of the theory to be an a posteriori attempt to salvage it after its central predictions fail. Proponents should abandon the "baby in the bathwater" argument in either a falsificatory or a confirmatory epistemological context.

Using a confirmatory approach, G&A nonrandomly select three of our 46 data sets for reanalysis, but instead of fitting the best model under least squares, they force a slope of −0.65 and interpret the explanatory power of their model based on coefficients of determination. Notably, one of the groups that they selected (tiger beetles) is inconsistent with two of their restrictive conditions, being a narrowly defined taxonomic group and comprising species that thermoregulate (Pearson and Vogler 2001, Dajoz 2002). They also select amphibians, but many of these also thermoregulate (Hutchinson and Dupré 1992). This illustrates the difficulty in understanding when the theory applies. Irrespectively, we repeated their approach for all 46 data sets, ignoring any nonlinearity following G&A but violating the postulate of linearity by Allen et al. (2002). The coefficients of determination of these tests were very low, with 27 being zero, and eight others being less than 0.30 (Table 1). Across all data sets, the r^2 values were substantially lower than the r^2 values from OLS fits (paired t test = −5.39; $P < 0.001$), despite low overall fits of temperature using either method (average $r^2_{\mathrm{G\&A}}$ =

TABLE 1. Coefficients of determination (r^2) for linear regressions of ln-transformed richness against rescaled temperature using ordinary least squares [OLS] vs. the "forced slope" method of Gillooly and Allen (2007) [G&A].

Group	Region	OLS	G&A
Amphibians	Afrotropics	0.216	0
Amphibians	Australia	0.003	0
Amphibians	Brazil	0.590	0
Amphibians	China	0.404	0.384
Amphibians	Europe	0.502	0.499
Amphibians	Iberia	0.010	0
Amphibians	North America	0.767	0.739
Angiosperms	China	0.353	0.059
Ants	Colorado/Nevada	0.053	0
Ants	New World	0.582	0.545
Blister beetles	North America	0.347	0.312
Bumble bees	global	0.165	0
Butterflies	Australia	0.030	0
Butterflies	California	0.306	0
Butterflies	western Palearctic	0.136	0
Butterflies (summer)	North America	0.606	0.261
Butterflies (winter)	North America	0.499	0
Dung beetles	Iberia/France	0.008	0
Dung beetles	western Palearctic	0.111	0
Eupelmid wasps	Palearctic	0.084	0
Grasshoppers	North America	0.284	0.016
Hawk moths	Mexico	0.221	0.209
Hawk moths	southeastern Asia	0.025	0
Orthoptera	Catalonia	0.101	0
Plants	California	0.192	0
Plants	Catalonia	0.044	0
Plants (exotic)	Great Britain	0.656	0.190
Plants (native)	Great Britain	0.519	0.490
Pteridophytes	Europe	0.208	0
Pteridophytes	Iberia	0.055	0
Reptiles	Brazil	0.014	0
Reptiles	China	0.383	0.381
Reptiles	Europe	0.607	0.588
Reptiles	Iberia	0.002	0
Reptiles	North America	0.838	0.620
Reptiles	southern Africa	0	0
Seed plants	Iberia	0.082	0
Snakes	Afrotropics	0.278	0
Tiger beetles	Australia	0.113	0.099
Tiger beetles	India	0	0
Tiger beetles	North America	0.560	0.544
Tiger beetles	northwestern South America	0.156	0.153
Trees	Europe	0.458	0.372
Trees	North America	0.588	0.584
Woody plants	Kenya	0.338	0
Woody plants	southern Africa	0.019	0.012

Note: Regressions were done across all values within each of the 46 data sets ignoring any nonlinearity in the data.

0.153; average $r^2_{\mathrm{OLS}} = 0.272$). Although we currently do not have other environmental predictors for all data sets, previous meta-analyses (Hawkins et al. 2003) indicate that r^2 values of other variables (derived from theories related to water–energy balance; e.g., O'Brien [2006]) have much greater statistical explanatory power. Moreover, recent modeling of geographic range overlap explicitly based on MTE generated results with lower explanatory power than those generated using alternative models (Rahbek et al. 2007).

If the confirmatory approach is to be used for testing MTE, and any positive relationship between temperature and diversity is "promising," evaluations will become mainly correlative, as with many competing theories. Therefore, model developers must clearly describe the unique predictions made by their model (Shipley 2000, Currie et al. 2004). This is essential for understanding diversity gradients, because the spatial structure of climatic variation on Earth causes nearly all theories developed to explain broadscale richness gradients to predict a positive correlation between richness and temperature, even when no causal link between them exists, such as in the "pure tropical conservatism" model (Wiens and Donoghue 2004).

CONCLUDING REMARKS

MTE can be viewed as the core of a research program. The hypothesis of Allen et al. (2002), together with the model(s) developed to test it, is one facet of the program. Their model(s) can be tested and rejected, but this does not necessarily challenge the core. As pointed out by Hawkins et al. (2007), our evaluation was restricted to the predictions of Allen et al. (2002) and Brown et al. (2004) for richness gradients and cannot be generalized to MTE as a whole (also see Latimer 2007). Even so, we contend that the tests by Hawkins et al. (2007) are as valid as proponents' tests and provide strong evidence against the model as a *general* explanation. Of course, it is difficult to know whether the failure of the model's predictions occurs at the postulate, hypothesis, or theory level. Incorporating additional variables (including spatial variation in average body size and abundance, as well including potential deviations from the energetic equivalence rule) might generate improved models that better fit the empirical data. Perhaps this could support the claim that MTE explains richness gradients, at least in part (see also Latimer 2007). But arguing that it might and showing to what extent it does are very different propositions.

LITERATURE CITED

Algar, A. C., J. T. Kerr, and D. J. Currie. 2007. A test of metabolic theory as the mechanism underlying broad-scale species-richness gradients. Global Ecology and Biogeography 16:170–178 [doi: 10.1111/j.1466-8238.2006.oo275x].

Allen, A. P., J. F. Gillooly, and J. H. Brown. 2002. Global biodiversity, biochemical kinetics, and the energetic-equivalence rule. Science 297:1545–1548.

Allen, A. P., J. F. Gillooly, V. M. Savage, and J. H. Brown. 2006. Kinetic effects of temperature on rates of genetic divergence and speciation. Proceedings of the National Academy of Sciences (USA) 103:9130–9135.

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

Currie, D. J., G. G. Mittelbach, H. V. Cornell, R. Field, J.-F. Guégan, B. A. Hawkins, D. M. Kaufman, J. T. Kerr, T. Oberdorff, E. O'Brien, and J. R. G. Turner. 2004. Predictions and tests of climate-based hypotheses of broad-scale variation in taxonomic richness. Ecology Letters 7:1121–1134.

Dajoz, R. 2002. Les coléoptères carabidés et ténébrionidés: écologie et biologie. Tec & Doc, Paris, France.

Field, R., E. M. O'Brien, and R. J. Whittaker. 2005. Global models for predicting woody plant richness from climate: development and evaluation. Ecology 86:2263–2277.

Ford, E. D. 2002. Scientific method for ecological research. Cambridge University Press, Cambridge, UK.

Gillooly, J. F., and A. P. Allen. 2007. Linking global patterns in biodiversity to evolutionary dynamics using metabolic theory. Ecology 88:1890–1894.

Graham, M. H. 2003. Confronting multicollinearity in ecological multiple regression. Ecology 84:2809–2815.

Hawkins, B. A., et al. 2007. A global evaluation of metabolic theory as an explanation for terrestrial species richness gradients. Ecology 88:1877–1888.

Hawkins, B. A., J. A. F. Diniz-Filho, C. A. Jaramillo, and S. A. Soeller. 2006. Post-Eocene climate change, niche conservatism, and the latitudinal diversity gradient of New World birds. Journal of Biogeography 33:770–780.

Hawkins, B. A., J. A. F. Diniz-Filho, C. A. Jaramillo, and S. A. Soeller. *In press*. Climate, niche conservatism, and the global bird diversity gradient. American Naturalist.

Hawkins, B. A., J. A. F. Diniz-Filho, and S. A. Soeller. 2005. Water links the historical and contemporary components of the Australian bird diversity gradient. Journal of Biogeography 32:1035–1042.

Hawkins, B. A., R. Field, H. V. Cornell, D. J. Currie, J. F. Guégan, D. M. Kaufman, J. T. Kerr, G. G. Mittelbach, T. Oberdorff, E. M. O'Brien, E. E. Porter, and J. R. G. Turner. 2003. Energy, water, and broad-scale geographic patterns of species richness. Ecology 84:3105–3117.

Hilborne, R., and M. Mangel. 1997. The ecological detective: confronting models with data. Princeton University Press, Princeton, New Jersey, USA.

Hutchinson, V. H., and R. K. Dupré. 1992. Thermoregulation. Pages 206–249 *in* M. E. Feder and W. W. Burggren, editors. Environmental physiology of the amphibians. University of Chicago Press, Chicago, Illinois, USA.

Latimer, A. M. 2007. Geography and resource limitation complicate metabolism-based predictions of species richness. Ecology 88:1895–1898.

Lavers, C., and R. Field. 2006. A resource-based conceptual model of plant diversity that reassesses causality in the productivity–diversity relationship. Global Ecology and Biogeography 15:213–224.

McGill, B. 2003. Strong and weak tests of macroecological theory. Oikos 102:679–685.

O'Brien, E. M. 2006. Biological relativity to water–energy dynamics. Journal of Biogeography 33:1868–1888.

Pearson, D. L., and A. P. Vogler. 2001. Tiger beetles: the evolution, ecology, and diversity of the cicindelids. Cornell University Press, Ithaca, New York, USA.

Rahbek, C., N. J. Gotelli, R. K. Colwell, G. L. Entsminger, T. F. L. V. B. Rangel, and G. R. Graves. 2007. Predicting continental-scale patterns of bird species richness with spatially explicit models. Proceedings of the Royal Society B 274:165–174 [doi: 10.1098/rspb.2006.3700].

Shipley, B. 2000. Cause and correlation in biology. Cambridge University Press, Cambridge, UK.

Whittaker, R. J., D. Nogués-Bravo, and M. B. Araújo. 2007. Geographic gradients of species richness: a test of the water–energy conjecture of Hawkins et al. (2003) using European data for five taxa. Global Ecology and Biogeography 16:76–89.

Whittaker, R. J., K. J. Willis, and R. Field. 2001. Scale and species richness: towards a general, hierarchical theory of species diversity. Journal of Biogeography 28:453–470.

Wiens, J. J., and M. J. Donoghue. 2004. Historical biogeography, ecology and species richness. Trends in Ecology and Evolution 19:639–644.

Willig, M. R., D. M. Kaufman, and R. D. Stevens. 2003. Latitudinal gradients of biodiversity: pattern, process, scale and synthesis. Annual Review of Ecology, Evolution and Systematics 34:273–309.