

Ecology, 88(8), 2007, pp. 1877–1888
 © 2007 by the Ecological Society of America

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-SANUDO,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 WILLIAMS¹⁷

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 Biología 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 Biología, 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. 2006a, 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) In-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 In-transformed richness and $1000/K$ should be -9.0 .

Brown et al. (2004) and Allen et al. (2007) subsequently used a different rescaling, $1/(kK)$, where k is the Boltzmann’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 ($1 \text{ eV} = 1.602 \times 10^{-19} \text{ J}$). 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 $2.5^\circ \times 2.5^\circ$ in the south and $2.5^\circ \times 5.0^\circ$ in the north, trees along an elevational gradient in Costa Rica, North American amphibians in $2.5^\circ \times 2.5^\circ$ and $2.5^\circ \times 5.0^\circ$ 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 $611\,000 \text{ km}^2$. 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	e'	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 (ν^*), 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}_w was calculated using the reciprocals of the squared standard errors ($1/SE^2$) 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.

Group	Region	Figure	Break point	Cool				Warm			
				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	Palaearctic	b'	40.8	0.79	<0.001	−0.67	−0.75	0.31	0.03	+0.32	+0.57
Butterflies	western Palaearctic	a'	41.1	0.38	0.07	−0.57	−0.92	0.62	0.008	+1.25	+1.59
Dung beetles	western Palaearctic	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 Palaearctic (Fig. 1B: b', data to the right of the

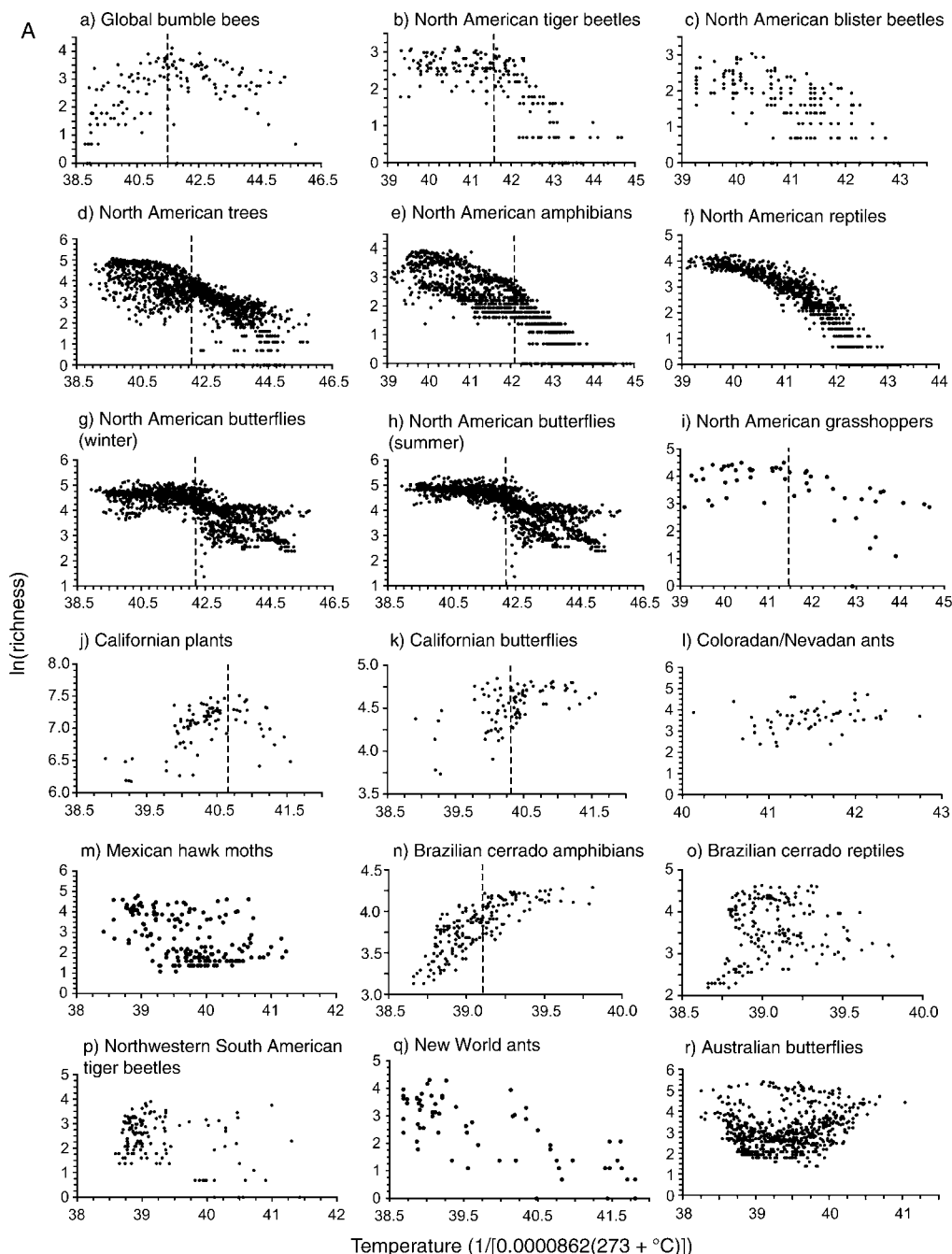


FIG. 1A–C. Scatterplots of the data sets included in the analysis; richness is the number of species. Dashed vertical lines are break points used to divide nonlinear data into linear components. Note that the temperature variable is a reciprocal; actual temperature decreases from left to right. “Fig. 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:

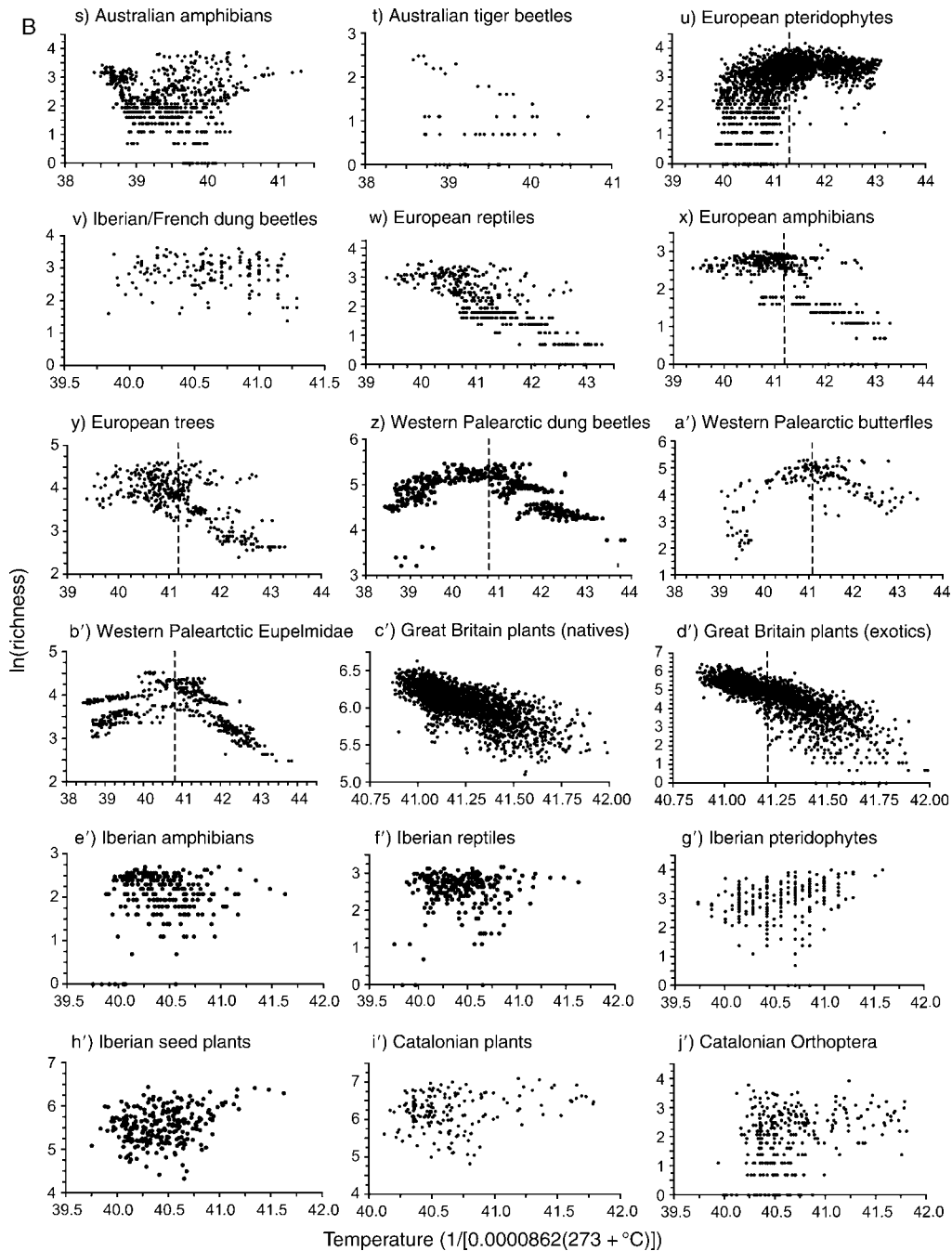


FIG. 1B. Continued.

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.

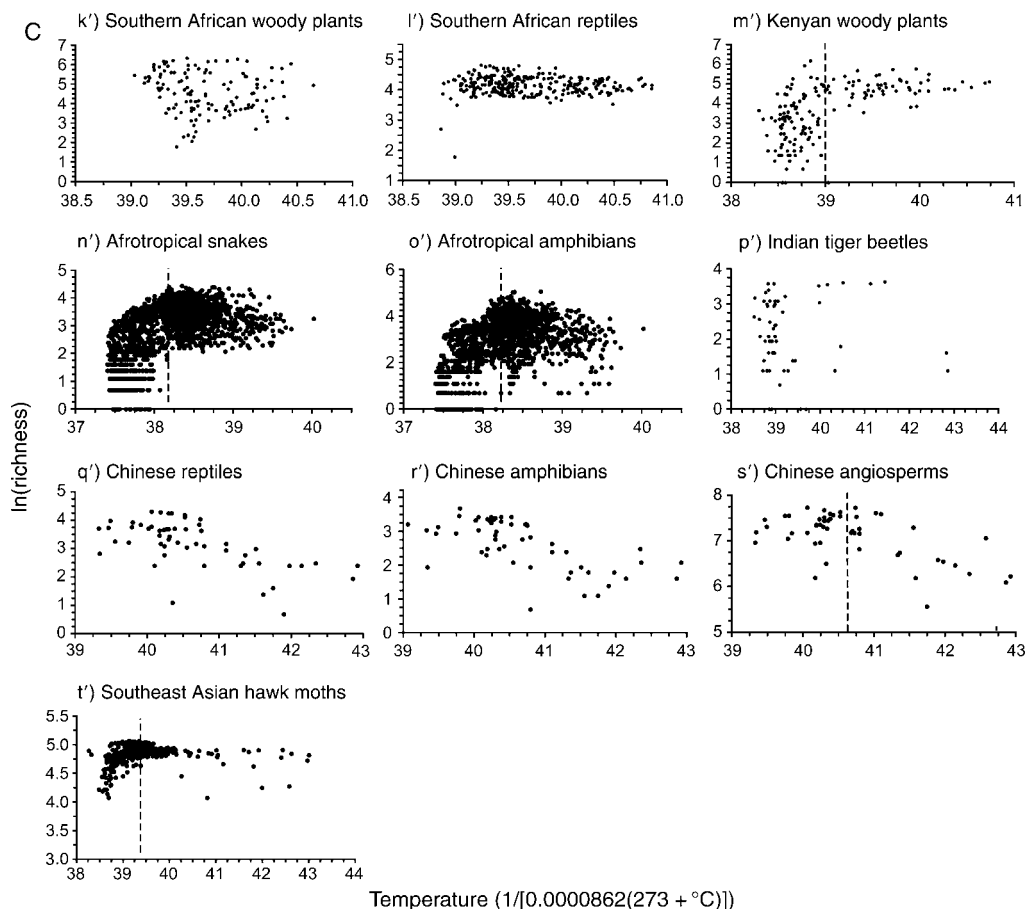


FIG. 1C. Continued.

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 \pm 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,

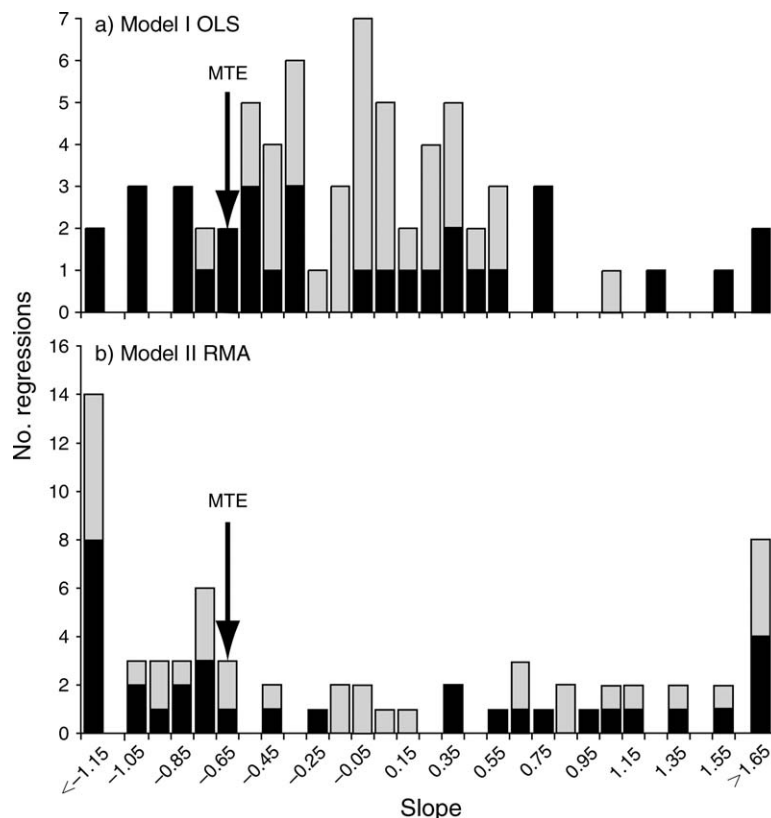


FIG. 2. Frequency distribution of slopes of all (a) OLS and (b) RMA regressions (see Table 1). The arrows identify the class containing the slope predicted by metabolic theory. Black bars represent statistically significant ($P < 0.05$) regressions, whereas nonsignificant regressions ($P > 0.05$) are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

ear 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 "broad-scale" 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.1466-8238.2006.00275x].
- 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* 423:639–642.
- 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. 2006a. 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. 2006b. 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. (2003a) 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).