

Ecology, 88(8), 2007, pp. 1895–1898
© 2007 by the Ecological Society of America

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

ANDREW M. LATIMER¹

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.

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

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

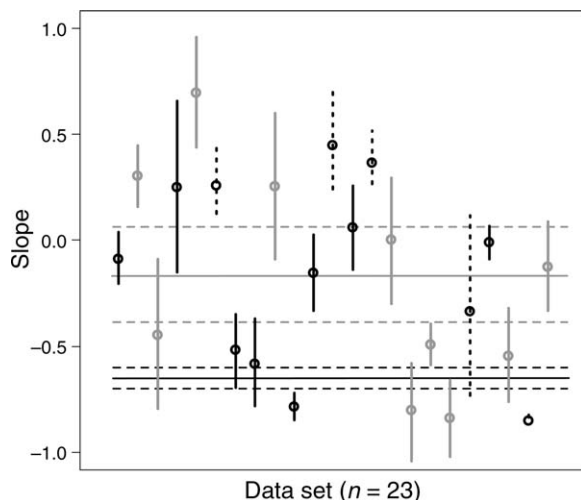


FIG. 1. Means and credible intervals for the slope parameters of 23 linear data sets. Each point represents the posterior mean of the slope for one data set in the hierarchical analysis, and the vertical lines span the 95% credible interval for the slope. The different line types indicate broad taxonomic 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

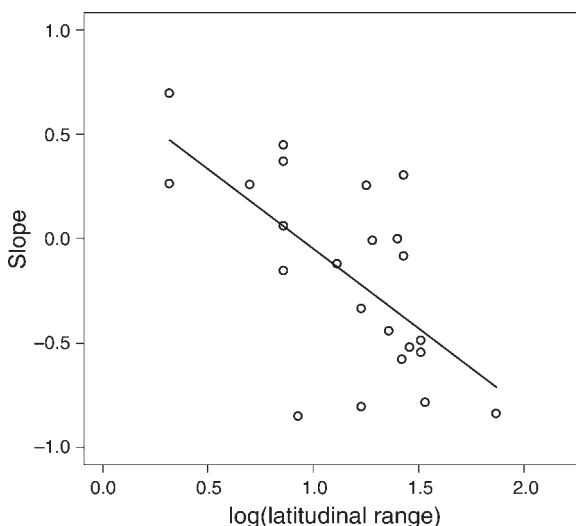


FIG. 2. Scatter plot of inferred slopes against log-transformed latitudinal extent (range in degrees) for the 23 linear data sets. The trend line is the least-squares linear fit through the points.

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 (Sturner 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 envi-

ronments, 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
© 2007 by the Ecological Society of America

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. SANDERS⁸

¹Department of Ecology and Evolutionary Biology, University of California, Irvine, California 92697 USA

²Departamento de Biologia Geral, ICB, Universidade Federal de Goiás, CP 131, 74.001-970, Goiânia, GO, Brazil

³Departamento de Biodiversidad y Biología Evolutiva, Museo Nacional de Ciencias Naturales (CSIC), 28006 Madrid, Spain

⁴School of Geography, University of Nottingham NG7 2RD UK

⁵Department of Biology, University of Ottawa, Ottawa, Ontario K1N 6N5 Canada

⁶Center for Macroecology, Institute of Biology, University of Copenhagen, DK-2100 Copenhagen, Denmark

⁷Departamento de Ecología, Universidad de Alcalá, 28871 Alcalá de Henares, Madrid, Spain

⁸Department 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.

⁹E-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 non-confirmatory 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 $r_{G\&A}^2 =$

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 $r^2_{OLS} = 0.272$). 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.1466-8238.2006.00275x].
- 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, 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].
- Shiple, 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.