

The mid-domain effect matters: simulation analyses of range-size distribution data from Mount Kinabalu, Borneo

John-Arvid Grytnes^{1,2}*, John H. Beaman^{3,4}, Tom S. Romdal² and Carsten Rahbek²

¹Department of Biology, University of Bergen, Bergen, Norway, ²Center for Macroecology, Institute of Biology, University of Copenhagen, Copenhagen, Denmark, ³The Herbarium, Royal Botanic Gardens, Kew, Surrey, UK and ⁴Department of Plant Biology, Michigan State University, East Lansing, MI, USA

ABSTRACT

Aim In simulation exercises, mid-domain peaks in species richness arise as a result of the random placement of modelled species ranges within simulated geometric constraints. This has been called the mid-domain effect (MDE). Where close correspondence is found between such simulations and empirical data, it is not possible to reject the hypothesis that empirical species richness patterns result from the MDE rather than being the outcome (wholly or largely) of other factors. To separate the influence of the MDE from other factors we therefore need to evaluate variables other than species richness. The distribution of range sizes gives different predictions between models including the MDE or not. Here, we produce predictions for species richness and distribution of range sizes from one model without the MDE and from two MDE models: a classical MDE model encompassing only species with their entire range within the domain (rangerestricted MDE), and a model encompassing all species with the theoretical midpoint within the domain (midpoint-restricted MDE). These predictions are compared with observations from the elevational pattern of range-size distributions and species richness of vascular plants.

Location Mount Kinabalu, Borneo.

Methods The data set analysed comprises more than 28,000 plant specimens with information on elevation. Species ranges are simulated with various assumptions for the three models, and the species simulated are subsequently subjected to a sampling that simulates the actual collection of species on Mount Kinabalu. The resulting pattern of species richness and species range-size distributions are compared with the observed pattern.

Results The comparison of simulated and observed patterns indicates that an underlying monotonically decreasing trend in species richness with elevation is essential to explain fully the observed pattern of richness and range size. When the underlying trend is accounted for, the MDE model that restricts the distributions of theoretical midpoints performs better than both the classical MDE model and the model that does not incorporate geometric constraints.

Main conclusions Of the three models evaluated here, the midpoint-restricted MDE model is found to be the best for explaining species richness and species range-size distributions on Mount Kinabalu.

Keywords

Altitudinal gradients, Borneo, geometric constraints, mid-domain effect, null models, range-size distribution, species richness.

*Correspondence: John-Arvid Grytnes, Department of Biology, University of Bergen, Allegaten 41, 5007 Bergen, Norway. E-mail: john.grytnes@bio.uib.no

INTRODUCTION

Numerous hypotheses have been proposed to explain geographical gradients in species richness. Those associated with spatial variation in climate are the oldest and remain the most-cited explanations for broad-scale patterns of species richness (Forster, 1778; Wallace, 1878; O'Brien, 1993; Rahbek, 1997; Hawkins et al., 2003; Currie et al., 2004). Other commonly cited factors that may influence species richness are area (Terborgh, 1973; Rosenzweig, 1995; Rahbek, 1997), and aspects of history (Wallace, 1878; Fischer, 1960; Jetz et al., 2004; Rahbek et al., 2007). The mid-domain effect (MDE) has recently been added to the list of possible models for predicting broad-scale species-richness patterns (Willig et al., 2003; Colwell et al., 2004; Pimm & Brown, 2004). This model is based on the random distribution of ranges within a restricted geographical domain (Colwell & Hurtt, 1994; Colwell & Lees, 2000; Colwell et al., 2004), resulting in a mid-domain peak or plateau in species richness without the need for any climatic or environmental factors for the placement of these ranges.

Many recent studies of broad-scale species richness have thus evaluated the potential role of the MDE by comparing model predictions with observed patterns. Colwell et al. (2004) reviewed 21 published MDE studies and concluded that a substantial signature of the MDE could be detected in natural patterns. This conclusion has, however, been criticized (Zapata et al., 2005). One of the reasons for this disagreement is perhaps that critics of the MDE tend to think about the matter as a dichotomy – MDE either explains all or nothing – whereas proponents of the MDE argue that the MDE should not have primacy before other factors, but should be studied as a potential factor along with other explanatory factors (Colwell et al., 2004, 2005; Lees & Colwell, 2007). In addition, although MDE predictions often correspond well to the observed species-richness pattern, this pattern may have other causes as well. Other factors, such as climate and area, may in many cases produce the same humped pattern as the MDE, and in these cases comparing predicted and observed species-richness patterns is not sufficient to separate the various hypotheses. The interpolation of species ranges between observed extremes may also create a species-richness pattern that may be confounded with the MDE predictions (Grytnes & Vetaas, 2002; Zapata et al., 2003). To separate the various factors we therefore need to look at variables other than the speciesrichness pattern. In addition to predicting a humped speciesrichness pattern, the MDE predicts that range sizes will have different distributions along the domain from what is predicted from models not incorporating the MDE (Grytnes, 2003; Arita, 2005). More specifically, the range-size distributions of the species that are found close to the domain limits will be different depending on whether the MDE is assumed or not, so that for models not incorporating the MDE the observed frequency distribution of range sizes (i.e. for the part of the range that is within the domain) will tend to have many more species with small ranges compared with models

Simulation models can be seen as ecologically controlled quasi-experiments in which some factors are held constant in order to isolate the effect of other potential processes on the observed pattern (Peck, 2004; Rangel et al., 2007). We will in this study use simple simulations to study the potential effect of the MDE along an altitudinal gradient of species richness. From the simulation models we make predictions on both species-richness patterns and species range-size distributions from models with and without the MDE, and we compare these predictions with empirical data covering the elevational distribution of vascular plants on a species-rich tropical mountain, Mount Kinabalu in Borneo. A data set describing the elevational distribution of vascular plant species on this mountain has been developed recently from herbarium data collected over the last one and a half centuries (Beaman, 2005); it includes more than 28,000 specimens with exact information on specimen elevation. This detailed information allows us to take into account the influence of sampling differences when comparing the observed and simulated patterns of species richness and range-size distribution.

METHODS

Study area and data

Mount Kinabalu is located at approximately 6°05' N, 116°35' E. Rising to 4094 m a.s.l., it is the highest mountain in Southeast Asia. The geology of the area is complex (Collenette, 1958; Jacobson, 1978), and the mountain is one of the youngest mountains on Earth, still rising at a rate of about 5 mm a year (Tain Choi, 1996). During the Pleistocene, the summit supported an ice cap of about 5 km². Deglaciation of the summit occurred c. 9200 years ago (Jacobson, 1978). The mountain has a humid tropical climate, and the mean annual temperature at sea level is 27.5°C, with a lapse rate of 0.55°C per 100 m (Kitayama, 1992; Kitayama et al., 1999). There are no simple trends in precipitation pattern with elevation (Kitayama, 1992; Kitayama & Aiba, 2002), as the cloud-zone moisture increases steeply between 1500 and 2000 m, where the latter elevation corresponds to the lower cloud level and precipitation is lower both above and below this elevation belt (Kitayama, 1992).

In connection with a recent enumeration of the flora on Mount Kinabalu (Parris *et al.*, 1992; Wood *et al.*, 1993; Beaman & Beaman, 1998; Beaman *et al.*, 2001; Beaman & Anderson, 2004), a data base of the present collections has been developed by Beaman and colleagues (Beaman, 2005). Details of the treatment of specimens from the data base prior to analyses are outlined in Grytnes & Beaman (2006). The total number of specimens with sufficient information to permit an analysis of distribution based on elevation is 28,458. For the purpose of describing the species-richness pattern we binned the data into 300-m elevational intervals. When plotting the frequencies of specimens along the elevational gradient at 100-m intervals, a peak was observed for each 300 m. This artefact is caused by the collectors' tendency to use round numbers, resulting in a peak for each thousand feet. The use of 300-m intervals produces a significantly smoother curve when plotting the frequencies of specimens along elevation.

Models

The different MDE models that have been proposed predict different distributions of range sizes in the domain. To account for this, two different MDE models are evaluated. The first suite of MDE models was proposed by Colwell & Hurtt (1994). A fundamental condition of the early MDE models is that they require entire ranges to be placed within the domain (Colwell et al., 2004). Grytnes & Vetaas (2002) developed a model that is based on potential midpoints and potential ranges, which can expand outside the domain. These two MDE models generate only slightly different predictions for species-richness patterns, but give clearly different predictions for range-size distributions, at least for those species that are found close to the domain boundaries (Grytnes, 2003). As a result of this, three distinct models are simulated. The first model (hereafter referred to as Model 0) is developed following the assumption that there are no hard boundaries. The two MDE models differ with respect to how the ranges are restricted within the domain. The first MDE model restricts the distribution of the entire ranges, so that, if a part of the range of a species is deemed to lie outside the domain, this species is not included in the simulations (Colwell & Hurtt, 1994; Colwell & Lees, 2000; Colwell et al., 2004). The second MDE model places restrictions on the distribution of the potential midpoint of a species. A potential or theoretical range is simulated for each species, and this theoretical range is not restricted and can expand beyond the domain limits. The part of the theoretical range that expands outside the domain is deleted, and the remaining 'observed range' remains within the domain (Grytnes & Vetaas, 2002). Because of this primary difference between the two models, we will subsequently refer to the latter model as the midpoint-restricted MDE (MrMDE) model and to the former as the range-restricted MDE (RrMDE) model. The RrMDE model has not been allocated any explicit ecological interpretation, but is perceived as a null model in which the observed ranges are randomly distributed along the domain (Colwell et al., 2004). For the MrMDE model, the potential midpoint for a species is the elevation at which it has its maximum abundance or evolutionary origin (Grytnes & Vetaas, 2002; Grytnes, 2003).

Details of the model simulations are as follows.

Model 0 assumes that the domain boundaries have no effect on species distributions. This model is simulated as described in Grytnes & Vetaas (2002; Model 0) and returns the same result as Model 1 of Colwell & Hurtt (1994). This model allows the potential midpoint to be placed far below or above the hard boundaries. As the maximum range size is set to 3500 m, we allow a potential midpoint to be found between 1750 m above the highest point and 1750 m below sea level. This corresponds to species having optima in warmer or colder climates than can be found in the domain (Grytnes & Vetaas, 2002). Species not reaching the domain are not included in further analyses.

The **RrMDE model** is simulated by excluding all species from Model 0 that extend beyond the hard boundaries, so that all species that have part of their ranges outside the domain are deleted. Our simulation corresponds to the fully stochastic model of Colwell & Hurtt (1994; their Model 2).

The **MrMDE model** is simulated in the same way as described in Grytnes & Vetaas (2002) and as the evolutionary or source-sink MDE model in Grytnes (2003). It is simulated by restricting the potential midpoint to be placed in the domain, i.e. between sea level and mountain top. No restrictions are put on the potential range, but the parts of the potential ranges that are found outside the domain are eliminated and the resulting simulated 'observed' ranges are thus completely within the domain.

For the RrMDE model, one should ideally use the observed distribution of species ranges as advocated in Colwell et al. (2004). However, despite the high number of specimens collected, the majority of the species ranges of the species in our data base have probably not been completely sampled. For the MrMDE model the observed ranges would not be suitable even if the ranges were estimated correctly, as we are not able to draw information regarding the potential ranges from observed data. Our simulations are thus based entirely on theoretical distributions of range sizes and midpoints for all models. Because we do not know the 'correct' theoretical range-size distribution, we tried several distributions. First, a uniform distribution was assumed and the theoretical range sizes were set to vary between 0 and 3500 m with equal probability for all range sizes. The maximum range size in the empirical data set was 3390 m. Second, lognormal distributions of the theoretical range-size distributions were assumed. For the lognormal distributions we assumed a mean value of seven (corresponding to 1097 m on a natural scale), with eight different standard deviations (0.2, 0.4, 0.6, 0.8, 1.0, 1.2, 1.4 and 1.6). Other mean values and standard deviations were tried, but simulations with other values deviated more from the empirical observations than the ones chosen and are not shown.

The summit of Mount Kinabalu is 4094 m a.s.l., but some of the earlier samples have been given an elevation of 13,500 ft, which corresponds to 4116 m. This latter elevation is here defined to be the upper hard boundary. No attempt has been made to correct any original measure. The lower boundary is set as sea level. The lowest point in the defined study area is around 200 m, but the distance to the sea is short (c. 40 km from the summit of Mount Kinabalu), and, as a result of this, sea level has been set as the lower boundary. Very few specimens have been collected below 300 m.

As pointed out by several authors, considering the MDE as the sole variable may lead to the erroneous conclusion that the MDE has no effect on species-richness patterns (Grvtnes & Vetaas, 2002; Colwell et al., 2004; Rahbek, 2005). We therefore investigated the effect of the MDE after accounting for a simple underlying trend in species richness as well as considering the MDE alone. This underlying trend is simulated for each of the three models by systematically varying the probability of placement of potential midpoints along elevation. An estimate of the underlying trend is found by using the uppermost 2100 m (seven uppermost 300-m intervals) in a generalized linear model with Poisson error distribution and logarithmic link. The resulting parameters from this regression are used to describe the probability distribution of potential midpoints of species along the elevational gradient. The probability distribution is then equal to $exp(9.15 + 0.001128 \times elevation)$. A uniform distribution of potential midpoints along elevation is assumed for models lacking an underlying trend. Various simple linear relationships for the potential midpoint probability distribution were also simulated, but these never gave a better fit to the empirical observations than a loglinear model and are not shown.

To compare the simulated models with the incompletely sampled empirical data, we performed sampling simulations, using detailed information regarding elevation for each specimen present in the empirical data set. For each collected specimen we randomly sampled one of the simulated species that 'exists' at the elevation of the collected specimen. This procedure was carried out for all 28,458 specimens in the data base, and thus the resulting elevational distribution of specimens from the simulations is identical to that observed from the data base.

The initial number of species simulated is different for each model because the range-size distribution of the simulated non-sampled models varies (cf. Grytnes, 2003), and this results in different probabilities of sampling species in the domain. Our aim was to have between 4000 and 4200 species in the final model after the simulation of sampling in order to have sufficient species to compare the range-size distributions from the simulations with the observed species range-size distribution.

Model evaluation

All the combinations of range-size distributions and probability distributions were tried in combination with the three main models. All the simulated models were evaluated initially by comparing the empirical observations with the simulated models, with special focus on range-size by plotting species ranges along the vertical axis against the rank of the species range on the horizontal axis. The range-size distribution that showed the best fit for each of the three MDE models was then analysed further in more detail and used in the comparison of the three models.

The most obvious parameter for comparing the MDE models is the species richness per elevational interval. This has

been the main focus for most studies that have considered MDE models (Lees et al., 1999; Jetz & Rahbek, 2001, 2002; Koleff & Gaston, 2001; Grytnes & Vetaas, 2002; Connolly et al., 2003; McCain, 2004, 2005). In our study, however, sampling intensity varies from 145 specimens at the highest elevational interval to over 8000 specimens at around 1500 m a.s.l. This large difference in sampling intensity will by randomness alone create an artificial pattern in species richness (the Pearson correlation coefficient between the number of specimens and number of species per elevational interval is 0.93), and it will also have a significant influence on the simulated richness pattern. If using 'observed' species richness from the different models, we will not be able to differentiate among the models because the 'observed' pattern is always heavily influenced by sampling intensity (the correlation between simulated 'observed' species richness for the different models generally results in a correlation coefficient between 0.95 and 0.99). We therefore compared the species-richness patterns by estimating richness per 300-m interval by rarefaction, as was done in Grytnes & Beaman (2006) to eliminate the influence of sampling on both the observed and simulated data sets (Gotelli & Colwell, 2001; Magurran, 2004). We randomly sampled 1000 specimens in the nine intervals that included more than 1000 collected specimens. We refrained from using correlation or regression because this gives an unnaturally high correlation for this type of modelling, suggesting the conclusion that these models explain all variation that possibly can be explained.

In addition to species richness, we also investigated rangesize distributions (Grytnes, 2003; Arita, 2005) by plotting species ranges along the vertical axis against the rank of the species range on the horizontal axis. The major difference in range size among the various models is expected to be at the domain boundaries (Grytnes, 2003). We therefore specifically investigated the range-size distributions of the species that are found in the lower and upper 1000 m and compared these with the range-size distributions of the species found only in the middle part of the domain.

RESULTS

The raw count of species found per 300-m elevational band is dependent on the number of collections made in each interval (Fig. 1). When rarefaction is used, empirical species richness increases moderately up to the interval between 900 m and 1200 m, and then decreases rapidly (Fig. 2). All models without an underlying trend in species richness perform poorly in predicting the empirical richness pattern, whether a uniform or a lognormal theoretical range-size distribution is used (Fig. 2; the results of the lognormal range-size distributions are shown in the Supplementary Material, Figs S1 & S2). When a loglinear underlying trend is assumed, the three models differ in how successfully they predict the empirical richness pattern. In theory, Model 0 should increase exponentially towards the lowermost elevation, but because rarefied species richness asymptotically approaches a ceiling of 1000 species (1000 specimens are used



Figure 1 Number of specimens (bars) and species richness (squares) of vascular plants per 300-m elevational interval on Mount Kinabalu. The total number of specimens is 28,458, and the total number of species is 3854. The midpoint of each interval is shown along the *x*-axis.

in the rarefactions), the estimated species-richness pattern creates the false impression that species richness levels off at the lowest elevation for this model (Fig. 2). The RrMDE model is the only model indicating a decrease in species richness at the lowermost elevation (Fig. 2), and the MrMDE model indicates that species richness stabilizes at the lowest interval studied here (Fig. 2). Towards higher elevations, the RrMDE model using a uniform range-size distribution decreases too slowly compared with the observed species richness (Fig. 2), whereas all RrMDE models using a lognormal range-size distribution follow the empirical pattern closely, even if the models overestimate richness at all intervals (Figs S1 & S2). All simulation models tend to overestimate species richness in most intervals.

The distribution of range sizes shows that there are relatively few species with large ranges, and that more than one-quarter of the species have an elevational range of 0 m (1013 out of 3854 species, of which 857 species were collected only once). Again, models with an underlying trend in species richness perform better than models assuming no underlying trend when comparing simulated patterns with observed patterns. We therefore focus on the results from the models that assume an underlying trend. The range-size distributions are presented for three subsets of species. First, the species found only in the middle of the domain (i.e. not found in the lower and upper 1000 m - these species are assumed not to be influenced by hard boundaries). Second, species found in the lower 1000 m; and third, species found in the upper 1000 m. For Model 0 and the MrMDE model, the uniform range-size distribution gives clearly better predictions than the lognormal distribution of range sizes. For the RrMDE model, some of the simulated models assuming a lognormal range-size distribution perform as well as, and, in some cases, even better than, the models assuming a uniform range-size distribution. Therefore, what we consider to be the best of these range-restricted models assuming a lognormal range-size distribution is presented together with the models assuming a uniform distribution in Fig. 3, and the other models assuming a lognormal range-size distribution are presented in Figs S1 & S2.

Looking at species found in the middle of the domain only, all the simulated models underestimate the number of species with very small ranges, and all four models shown give very similar predictions for these mid-domain species (Fig. 3, top row). The empirical range-size distributions for species found in the lower 1000 m are somewhat different from those found only in the middle of the domain, as relatively more species have ranges between 500 and 0 m (Fig. 3, middle row). Model 0 can now be seen to predict more species with small ranges compared with observations (Fig. 3, first column). Considering that this model had fewer species with small ranges when considering the mid-domain species, it is obvious that Model 0 is not capturing the difference in range-size distribution between these two species groups. Much the same trend is obvious for the RrMDE model that assumes a lognormal range-size distribution when



Figure 2 Rarefied species-richness pattern for the intervals from 300 to 2700 m compared with the simulated rarefied pattern for each of the three models assuming a uniform range-size distribution. The rarefaction is based on 1000 specimens for each 300-m interval. Black squares represent observed species richness. One hundred simulations are shown for each model. Light grey lines represent models without an underlying trend, and dark grey lines represent models with an underlying trend in species richness.



Figure 3 Plots of species range size versus rank of the range size for the empirical and simulated data. The simulations shown here are based on a uniform distribution of range sizes, except for the rightmost column, where a lognormal distribution is used with mean = 7 and standard deviation = 1. The black line is the pattern from the species in the empirical data set. The grey lines represent 100 simulations for each model. The panels in the top row represent the species that are found only between 1000 m and 3000 m. The panels in the middle row represent species that are found in the lowest 1000 m. The panels in the bottom row represent species that are found at 3000 m and upwards. The lines presented are a random draw of an appropriate number of species (without replacement) from each simulation. The number of species drawn is decided by the simulation that had least number of species in each category.

comparing species found in the middle of the domain only with species that occur in the lower 1000 m (Fig. 3, fourth column). The RrMDE model that assumes a uniform rangesize distribution is even less successful in capturing the observed change in range-size distribution between middomain species and species from the lowest 1000 m (Fig. 3, third column). The MrMDE model with an underlying trend of species richness is the model that adjusts best to the shape of the observed range sizes for the lowermost 1000 m compared with the range-size distribution for species found only in the middle of the domain (Fig. 3, second column). None of the models assuming a lognormal range-size distribution captures the observed change in range-size distribution between mid-domain species and species found in the lower 1000 m better than those shown in Fig. 3.

The observed range-size distribution for the species found in the uppermost 1000 m is different from the range-size

distribution of species found in the middle of the domain or in the lower part of the domain (Fig. 3, lower row). Here, the opposite pattern of what was found for the species occurring in the lowest 1000 m is observed; that is, there are relatively more species with large to intermediate range sizes. All models with a uniform range-size distribution perform almost equally, and the models assuming an underlying trend in species richness tend to predict too many species with large ranges compared with the empirical observations. As can be seen in Fig. 3, the variance in the 100 simulations is relatively large, so it is difficult to separate the models. However, it appears that Model 0 performs slightly better than the MrMDE model, which in turn performs slightly better than the RrMDE model. For the RrMDE model with lognormal range-size distribution, the number of species with intermediate to small range sizes is underestimated.

DISCUSSION

Species-richness pattern

The present study demonstrates that a model based on the MDE alone is not successful in predicting the observed species-richness pattern; however, when a simple underlying trend is assumed, the MDE models have an improved fit with the observed patterns in richness and range-size distributions. A study of elevational patterns in vascular plant species richness from Nepal also concluded that an MDE model, together with an underlying trend in species richness, explains the observed pattern far more effectively than either of the two variables alone (Grytnes & Vetaas, 2002), and a similar conclusion was reached by Rangel & Diniz-Filho (2005) and Storch et al. (2006). This demonstrates the importance of considering several factors simultaneously when investigating broad-scale species-richness patterns and of not prematurely discarding theories that are not able to explain species-richness patterns when used as a sole predictor (Willig et al., 2003; Colwell et al., 2004; Rahbek, 2005). In the following we will focus on models with an underlying gradient in species richness.

The simulation of species richness from the RrMDE model and the MrMDE model reveals a levelling off, or a decrease, in species richness towards the lower elevations, in accordance with the observed pattern of species richness. In this study, an exponential decrease in species richness with elevation is assumed. This may be incorrect, but the underlying trend in species richness is probably not steeper than this, at least not in the elevational intervals where the rarefactions are applied. The RrMDE model that assumes a uniform range-size distribution overestimates species richness at medium to high elevations (Fig. 2). This suggests that an even steeper gradient in underlying species richness should be used for this model in order to improve the correlation with the observed pattern of species richness, and we conclude that this model is probably not optimal for predicting species richness on Mount Kinabalu. In contrast, all the RrMDE models assuming a lognormal distribution of range sizes consistently overestimate the rarefied richness for all intervals, but the simulated pattern follows the empirical pattern closely. The MrMDE model overestimates species richness at the lowest elevations relative to the other elevation intervals, and a more modest increase in underlying species richness at the lowest elevations would give a perfect fit between observed and simulated species-richness trends.

There are two reasons why all simulated models tend to overestimate species richness at all elevation intervals in our study. First, the number of species simulated is deliberately exaggerated in order to ensure that sufficient species remain to permit the creation of rank-range diagrams for all species. For simulations performed here, it is not clear exactly how many species will remain after carrying out the simulated sampling. To ensure that our simulated models have at least as many sampled species as the observed data set (3854), we aimed for a total richness slightly in excess of 4000 species for each permutation. Most of the permutations thus have between 4000 and 4200 species. This level of species richness will result in a higher rarefied species richness. Second, although sampling is assumed to be random, this may not be true in certain respects. Collectors have a tendency to sample species with small ranges more often than would be expected from a random collection of individuals. The consequence of this will be that species that have large range sizes in the simulations will be 'over-sampled' in the simulations compared with the empirical data set. Relatively more sampling of large-ranged species (as in the simulations) would lead to an increased diversity per elevational interval if total richness on the whole mountain is kept constant. This would also result in an underestimation of small ranges in the simulated data when compared with the empirical data.

Range-size distribution

The observed range-size distribution shows that relatively many species have a small range size, in accordance with what has been documented previously (Brown et al., 1996; Hubbell, 2001; Gaston, 2003). Grytnes (2003) found that, when differentiating among MDE models (as well as models not including geometric constraints), range-size distribution will vary among the models, in particular near the domain limits. This is also evident from this study, as the difference between the simulated models here is most apparent when comparing the range-size distribution of species that are found close to the domain limit with the range-size distribution of species found only in the middle of the domain. Assuming a uniform range-size distribution and a complete sampling of all species in the domain, all models presented here will result in more species with narrow ranges than with large ranges for species found in the middle of the domain. Model 0 will have approximately the same distribution of range sizes for the species found in the outer part of the domain as for those found only in the middle of the domain. The RrMDE model will have an approximately uniform distribution of range sizes for the species found close to the domain limit, whereas the MrMDE model will have more species with intermediate range sizes than with small or large range sizes for the species found close to or at the domain (see also Grytnes, 2003). Comparing the upper and middle rows of Fig. 3, it is evident that the empirical pattern indicates that there are more species with ranges of around 500 m and smaller in the lower part than in the middle of the domain. After simulating the sampling, this is the same as found for the MrMDE model, whereas Model 0 and the RrMDE model assuming a lognormal range-size distribution show no indication of more species with range sizes around 500 m, but for both models there are predicted to be fewer small-ranged species relative to what was predicted for the species found in the middle of the domain. The RrMDE model assuming a uniform range-size distribution has too

many large-ranged species in the lowest 1000 m compared with the observed pattern. This means that the MrMDE model is the only model that is able to capture the observed empirical difference in range-size distribution between species found in the middle of the domain only and species found in the lowest 1000 m. However, the MrMDE model underestimates the number of small-ranged species, especially those with a range size equal to zero. Bearing in mind the possible sample effect mentioned above (collectors are more likely to sample species with a small range than species with a large range in the empirical data set), it is apparent that this model may actually be very close to being able to predict the rangesize distribution accurately, and clearly better than both Model 0 and the RrMDE model.

MDE or not?

When comparing predictions from theoretical models with observed data it will in most cases be possible to alter the parameters in the theoretical model to achieve a better correlation between the theoretical predictions and the observed data. All of the models simulated in this study could have been made more successful by making more detailed distributions of range size along elevation, by assuming more complex relationships between underlying richness and elevation, and by accounting for the biases in sampling (e.g. higher probability of sampling rare species relative to true abundance in the area). With the simple initial assumptions incorporated in the simulation models made here (uniform or lognormal range-size distribution), it appears that an exponential decrease in potential midpoint placement with elevation together with the MrMDE model produces the best fit to the observed pattern in species richness and range-size distribution. By assuming a modest decrease of species richness with elevation at the lower part of the gradient, the simulated species richness would show a perfect fit to the observed pattern. Trying to separate the three models detailed here, it is important to note that the MrMDE model was the only model able to capture the changes in the distributions found for the species occurring in the lowest 1000 m when compared with the distribution of range sizes for species found in the middle of the domain only. This comparison of species ranges in different parts of the domain gives the most consistent differences between these three models, independent of the parameters used for the models, and can therefore be seen as the strongest test of the models. A good fit between simulated and observed species richness can be found for Model 0 by assuming no underlying trends in species richness at the lowermost elevations. Range-size distributions for this model, however, were no better than for the MrMDE model, and this model did not capture the difference for the species appearing in the lower elevations. We therefore conclude that, with the simple assumptions made for our simulations, we need the MrMDE model, as well as environmental factors that create the underlying trend, to predict the observed speciesrichness and species range-size patterns on Mount Kinabalu.

ACKNOWLEDGEMENTS

We would like to thank H.J.B. Birks, V. Vandvik, and E. Heegaard for valuable comments on earlier versions of the manuscript, and B. H. Ingvartsen for help with the figures. J.A.G. received financial support from Norwegian Research Council, and J.H.B. was supported by grants from the US National Science Foundation.

REFERENCES

- Arita, H.T. (2005) Range size in mid-domain models of species diversity. *Journal of Theoretical Biology*, 232, 119–126.
- Beaman, J.H. (2005) Mount Kinabalu: hotspot of plant diversity in Borneo. *Biologiske Skrifter*, 55, 103–127.
- Beaman, J.H. & Anderson, C. (2004) The plants of Mount Kinabalu. 5. Dicotyledon families Magnoliaceae to Winteraceae. Natural History Publications (Borneo), Kota Kinabalu, in association with Royal Botanic Gardens, Kew.
- Beaman, J.H. & Beaman, R.S. (1998) The plants of Mount Kinabalu. 3. Gymnosperms and non-orchid monocotyledons. Natural History Publications (Borneo), Kota Kinabalu, in association with Royal Botanic Gardens, Kew.
- Beaman, J.H., Anderson, C. & Beaman, R.S. (2001) The plants of Mount Kinabalu. 4. Dicotyledon families Acanthaceae to Lythraceae. Natural History Publications (Borneo), Kota Kinabalu, in association with Royal Botanic Gardens, Kew.
- Brown, J.H., Stevens, G.C. & Kaufman, D.M. (1996) The geographic range: size, shape, boundaries, and internal structure. *Annual Review of Ecology and Systematics*, 27, 597–623.
- Collenette, P. (1958) *The geology and mineral resources of the Jesselton-Kinabalu area, North Borneo.* Memoir 6, Geological Survey Department, British Territories in Borneo, Kuching, Sarawak.
- Colwell, R.K. & Hurtt, G.C. (1994) Nonbiological gradients in species richness and a spurious Rapoport effect. *The American Naturalist*, **144**, 570–595.
- Colwell, R.K. & Lees, D.C. (2000) The mid-domain effect: geometric constraints on the geography of species richness. *Trends in Ecology and Evolution*, **15**, 70–76.
- Colwell, R.K., Rahbek, C. & Gotelli, N.J. (2004) The middomain effect and species richness patterns: what have we learned so far? *The American Naturalist*, **163**, E1–E23.
- Colwell, R.K., Rahbek, C. & Gotelli, N.J. (2005) The middomain effect: there's a baby in the bathwater. *The American Naturalist*, **166**, E149–E154.
- Connolly, S.R., Bellwood, D.R. & Hughes, T.P. (2003) Indo-Pacific biodiversity of coral reefs: deviations from a middomain model. *Ecology*, **84**, 2178–2190.
- Currie, D.J., Mittelbach, G.C., Cornell, H.V., Field, R., Guegan, J.F., Hawkins, B.A., Kaufman, D.M., Kerr, J.T., Oberdorff, T., O'Brien, E. & Turner, J.R.G. (2004) Predictions and tests of climate-based hypotheses of broad-scale variation in taxonomic richness. *Ecology Letters*, **7**, 1121–1134.
- Fischer, A.G. (1960) Latitudinal variation in organic diversity. *Evolution*, **14**, 64–81.

- Forster, J.R. (1778) Observations made during a voyage round the world, on physical geography, natural history, and ethic philosophy. G. Robinson, London.
- Gaston, K.J. (2003) The structure and dynamics of geographical ranges. Oxford University Press, Oxford.
- Gotelli, N.J. & Colwell, R.K. (2001) Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. *Ecology Letters*, **4**, 379–391.
- Grytnes, J.A. (2003) Ecological interpretations of the middomain effect. *Ecology Letters*, **6**, 883–888.
- Grytnes, J.A. & Beaman, J.H. (2006) Elevational species richness patterns for vascular plants on Mount Kinabalu, Borneo. *Journal of Biogeography*, **33**, 1838–1849.
- Grytnes, J.A. & Vetaas, O.R. (2002) Species richness and altitude: a comparison between simulation models and interpolated plant species richness along the Himalayan altitudinal gradient, Nepal. *The American Naturalist*, **159**, 294–304.
- Hawkins, B.A., Field, R., Cornell, H.V., Currier, D.J., Guegan, J.F., Kaufman, D.M., Kerr, J.T., Mittelbach, G.G., Oberdorff, T., O'Brien, E.M., Porter, E.E. & Turner, J.R.G. (2003) Energy, water, and broad-scale geographic patterns of species richness. *Ecology*, 84, 3105–3117.
- Hubbell, S.P. (2001) The unified neutral theory of biodiversity and biogeography. Princeton University Press, Princeton, NJ.
- Jacobson, G. (1978) Geology. *Kinabalu: summit of Borneo* (ed. by M. Luping, C. Wen and E.R. Dingley), pp. 101–110. Sabah Society Monograph, Kota Kinabalu.
- Jetz, W. & Rahbek, C. (2001) Geometric constraints explain much of the species richness patterns in African birds. *Proceedings of the National Academy of Sciences USA*, **98**, 5661–5666.
- Jetz, W. & Rahbek, C. (2002) Geographic range size and determinants of avian species richness. *Science*, 297, 1548–1551.
- Jetz, W., Rahbek, C. & Colwell, R.K. (2004) The coincidence of rarity and richness and the potential signature of history in centres of endemism. *Ecology Letters*, **7**, 1180–1191.
- Kitayama, K. (1992) An altitudinal transect study of the vegetation on Mount Kinabalu, Borneo. *Vegetatio*, **102**, 149– 171.
- Kitayama, K. & Aiba, S. (2002) Ecosystem structure and productivity of tropical rain forests along altitudinal gradients with contrasting soil phosphorous pools on Mount Kinabalu, Borneo. *Journal of Ecology*, **90**, 37–51.
- Kitayama, K., Lakim, M. & Wahab, M.Z. (1999) Climate profile of Mount Kinabalu during late 1995–early 1998 with special reference to the 1998 drought. *Sabah Parks Nature Journal*, **2**, 85–100.
- Koleff, P. & Gaston, K.J. (2001) Latitudinal gradients in diversity: real patterns and random models. *Ecography*, 24, 341–351.
- Lees, D.C. & Colwell, R.K. (2007) A strong Madagascan rainforest MDE and no equatorward increase in species richness: re-analysis of "The missing Madagascan mid-domain effect", by Kerr J.T., Perring M. & Currie D.J. (Ecology Letters 9:149–159, 2006). *Ecology Letters*, **10**, E4–E8.

- Lees, D.C., Kremen, C. & Andriamampianina, L. (1999) A null model for species richness gradients: bounded range overlap of butterflies and other rainforest endemics in Madagascar. *Biological Journal of the Linnean Society*, **67**, 529–584.
- Magurran, A.E. (2004) *Measuring biological diversity*. Blackwell Science Ltd, Oxford.
- McCain, C.M. (2004) The mid-domain effect applied to elevational gradients: species richness of small mammals in Costa Rica. *Journal of Biogeography*, **31**, 19–31.
- McCain, C.M. (2005) Elevational gradients in diversity of small mammals. *Ecology*, **86**, 366–372.
- O'Brien, E.M. (1993) Climatic gradients in woody plant species richness: towards an explanation based on an analysis of southern Africa's woody flora. *Journal of Biogeography*, **20**, 181–198.
- Parris, B.S., Beaman, R.S. & Beaman, J.H. (1992) The plants of Mount Kinabalu. 1. Ferns and fern allies. Royal Botanic Gardens, Kew.
- Peck, S.L. (2004) Simulation as experiment: a philosophical reassessment for biological modeling. *Trends in Ecology and Evolution*, **19**, 530–534.
- Pimm, S.L. & Brown, J.H. (2004) Domains of diversity. *Science*, **304**, 831–833.
- Rahbek, C. (1997) The relationship among area, elevation, and regional species richness in Neotropical birds. *The American Naturalist*, **149**, 875–902.
- Rahbek, C. (2005) The role of spatial scale and the perception of large-scale species-richness patterns. *Ecology Letters*, 8, 224–239.
- Rahbek, C., Gotelli, N.J., Colwell, R.K., Entsminger, G.L., Rangel, T.H.L.V.B. & Graves, G.R. (2007) Predicting continental-scale patterns of bird species richness with spatially explicit models. *Proceedings of the Royal Society B: Biological Sciences*, 274, 165–174.
- Rangel, T.H.L.V.B. & Diniz-Filho, A.F. (2005) An evolutionary tolerance model explaining spatial patterns in species richness under environmental gradients and geometric constraints. *Ecography*, 28, 253–263.
- Rangel, T.H.L.V.B., Diniz-Filho, A.F. & Colwell, R.K. (2007) Species richness and evolutionary niche dynamics: a spatial pattern-oriented simulation experiment. *The American Naturalist*, **170**, 602–616.
- Rosenzweig, M.L. (1995) Species diversity in space and time. Cambridge University Press, Cambridge.
- Storch, D., Davies, R.G., Zajiček, S., Orme, C.D.L., Olson, V., Thomas, G.H., Ding, T.-S., Rasmussen, P.C., Ridgley, R.S., Bennet, P.M., Blackburn, T.M., Owens, I.P.F. & Gaston, K.J. (2006) Energy, range dynamics and global species richness patterns: reconciling mid-domain effects and environmental determinants of avian diversity. *Ecology Letters*, **9**, 1308– 1320.
- Tain Choi, D.L. (1996) Geology of Kinabalu. Kinabalu: summit of Borneo, a revised and expanded edition (ed. by K.M. Wong and A. Phillipps), pp. 19–29. The Sabah Society in association with Sabah Parks, Kota Kinabalu.

- Terborgh, J. (1973) On the notion of favorableness in plant ecology. *The American Naturalist*, **107**, 481–501.
- Wallace, A.R. (1878) *Tropical nature and other essays*. Macmillan, New York.
- Willig, M.R., Kaufman, D.M. & Stevens, R.D. (2003) Latitudinal gradients of biodiversity: pattern, process, scale and synthesis. *Annual Review of Ecology, Evolution, and Systematics*, **34**, 273–309.
- Wood, J.J., Beaman, R.S. & Beaman, J.H. (1993) The plants of Mount Kinabalu. 2. Orchids. Royal Botanic Gardens, Kew.
- Zapata, F.A., Gaston, K.J. & Chown, S.L. (2003) Mid-domain models of species richness gradients: assumptions, methods and evidence. *Journal of Animal Ecology*, **72**, 677–690.
- Zapata, F.A., Gaston, K.J. & Chown, S.L. (2005) The middomain effect revisited. *The American Naturalist*, **166**, E144– E148.

SUPPLEMENTARY MATERIAL

The following supplementary material is available for this article:

Figure S1 Plots of rarefied species-richness pattern for the empirical and simulated data assuming various lognormal range-size distributions.

Figure S2 Plots of species range size versus rank of the range size for the empirical and simulated data assuming various lognormal range-size distributions.

This material is available as part of the online article from: http://www.blackwell-synergy.com/doi/abs/10.1111/j.1365-2699. 2008.01952.x (This link will take you to the article abstract). Please note: Blackwell Publishing is not responsible for the content or functionality of any supplementary materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

BIOSKETCHES

John-Arvid Grytnes is a postdoctoral researcher at the University of Bergen. His research interests are all aspects of species-richness patterns at different spatial scales. Current projects focus mostly on elevational richness patterns.

John H. Beaman is an honorary research fellow in the Herbarium at the Royal Botanic Gardens, Kew. With a number of collaborators he recently completed an inventory of the flora of Mount Kinabalu, a project extending over 22 years and listing about 5000 species.

Tom S. Romdal is a postdoctoral fellow at the Center for Macroecology at the University of Copenhagen. He works with explanations for macroecological patterns, mainly variation in species richness with latitude and elevation.

Carsten Rahbek is a professor at the University of Copenhagen and the director of the Center for Macroecology. His main research interests are patterns of species distribution, species range sizes, species assemblages and species richness, and what determines such patterns.

Editor: Lawrence Heaney