Diversity and Distributions, (Diversity Distrib.) (2011) **17**, 191–200

BIODIVERSITY RESEARCH

Funding begets biodiversity

Antje Ahrends[1,2,*], Neil D. Burgess[3,4,5], Roy E. Gereau[6], Rob Marchant[1], Mark T. Bulling[7,8], Jon C. Lovett[9], Philip J. Platts[1], Victoria Wilkins Kindemba[10], Nisha Owen[10], Eibleis Fanning[10] and Carsten Rahbek[5]

[1]*Environment Department, University of York, York YO105DD, UK,* [2]*Botanic Garden Edinburgh, 20A Inverleith Row, Edinburgh EH3 5LR, UK,* [3]*Zoology Department, University of Cambridge, Cambridge CB2 3EJ, UK,* [4]*WWF US, 1250 24th St NW, Washington, DC, USA,* [5]*Center for Macroecology, Evolution and Climate, Department of Biology, University of Copenhagen, Universitetsparken 15, DK-2100 Copenhagen Ø, Denmark,* [6]*Missouri Botanical Garden, PO Box 299, St. Louis, MO 63166-0299, USA,* [7]*Oceanlab, University of Aberdeen, Newburgh, Aberdeenshire, Aberdeen AB41 6AA, UK,* [8]*Department of Biology, Forensics and Sport, University of Derby, Derby DE22 1GB, UK,* [9]*Twente Centre for Studies in Technology and Sustainable Development, University of Twente, PO Box 217, 7500 Enschede, The Netherlands,* [10]*The Society for Environmental Exploration (Frontier), 50-52 Rivington Street, London EC2A 3QP, UK*

*Correspondence: Antje Ahrends, Royal Botanic Garden Edinburgh, 20A Inverleith Row, Edinburgh EH3 5LR, UK.
E-mail: aahrends@rbge.ac.uk

ABSTRACT

Aim Effective conservation of biodiversity relies on an unbiased knowledge of its distribution. Conservation priority assessments are typically based on the levels of species richness, endemism and threat. Areas identified as important receive the majority of conservation investments, often facilitating further research that results in more species discoveries. Here, we test whether there is circularity between funding and perceived biodiversity, which may reinforce the conservation status of areas already perceived to be important while other areas with less initial funding may remain overlooked.

Location Eastern Arc Mountains, Tanzania.

Methods We analysed time series data (1980–2007) of funding (n = 134 projects) and plant species records (n = 75,631) from a newly compiled database. Perceived plant diversity, over three decades, is regressed against funding and environmental factors, and variances decomposed in partial regressions. Cross-correlations are used to assess whether perceived biodiversity drives funding or *vice versa*.

Results Funding explained 65% of variation in perceived biodiversity patterns – six times more variation than accounted for by 34 candidate environmental factors. Cross-correlation analysis showed that funding is likely to be driving conservation priorities and not *vice versa*. It was also apparent that investment itself may trigger further investments as a result of reduced start-up costs for new projects in areas where infrastructure already exists. It is therefore difficult to establish whether funding, perceived biodiversity, or both drive further funding. However, in all cases, the results suggest that regional assessments of biodiversity conservation importance may be biased by investment. Funding effects might also confound studies on mechanisms of species richness patterns.

Main conclusions Continued biodiversity loss commands urgent conservation action even if our knowledge of its whereabouts is incomplete; however, by concentrating inventory funds in areas already perceived as important in terms of biodiversity and/or where start-up costs are lower, we risk losing other areas of underestimated or unknown value.

Keywords
Biodiversity inventories, conservation funding, conservation priorities, databiases, endemism, species richness.

INTRODUCTION

Biodiversity values are widely ranked according to species richness and the prevalence of threatened and endemic taxa (Brooks *et al.*, 2006). In conjunction with parameters such as levels of threat, these diversity metrics underpin several schemes that identify global (Stattersfield *et al.*, 1998; Olson & Dinerstein, 2002; Mittermeier *et al.*, 2005) and regional

DOI: 10.1111/j.1472-4642.2010.00737.x
http://wileyonlinelibrary.com/journal/ddi

Diversity and Distributions A Journal of Conservation Biogeography

(Fishpool & Evans, 2001; Eken *et al.*, 2004; Plantlife International, 2004) conservation priorities. Once an area is considered to be of high conservation priority, it typically attracts funding for further research and conservation. Priority locations such as Biodiversity Hotspots, Global 200 and Key Biodiversity Areas receive the majority of investments made available by global conservation funds and organizations (e.g. Dalton, 2000). Because intensification of research in an area is likely to result in the discovery of more species (Nelson *et al.*, 1990; Reddy & Davalos, 2003; Kier *et al.*, 2005; Soria-Auza & Kessler, 2008), including threatened and endemic species, the priority status of that area may be strengthened in a circular fashion (Küper *et al.*, 2004). Meanwhile, areas that have received little or no initial funding may remain perpetually overlooked. Thus, funding may beget perceived biodiversity importance and bias our understanding of conservation priorities.

Plant species richness is broadly related to environmental conditions, including levels of both anthropogenic and natural disturbance (O'Brien *et al.*, 2000; Hawkins *et al.*, 2003). Predictors for levels of endemism and the number of threatened species are less clear; however, both of these variables tend to increase with species richness (Jetz *et al.*, 2004). High levels of endemism have also been attributed to past climate configurations (Taplin & Lovett, 2003; Jetz *et al.*, 2004). We might therefore expect that environmental factors, including levels of disturbance, would be better predictors for perceived species diversity patterns than funding or survey effort. If inventory funding would emerge as an important predictor for perceived biodiversity, a number of explanations are possible: (1) biodiversity is driving funding; (2) inventory funding is driving perceived biodiversity, i.e. biasing our understanding of the distribution of biodiversity; (3) funding is driving perceived biodiversity which in turn drives funding; or (4) funding is driving perceived biodiversity and is also circularly associated with itself (for example, because areas with established project infra-structures have reduced costs for future investments). We tested these hypotheses in the Eastern Arc Mountains (EAM) of Tanzania (Fig. 3a), an area of outstanding biodiversity value. This range of ancient tropical mountains was a suitable test candidate as it is one of the most important sites for conservation globally (Stattersfield *et al.*, 1998; Olson & Dinerstein, 2002; Mittermeier *et al.*, 2005) and is one of the better studied global conservation priority areas, with botanical exploration going back over 130 years.

METHODS

Study area

The EAM are a chain of 13 ancient crystalline mountain blocs composed of heavily metamorphosed Precambrian basement rock and estimated to have been uplifted in the Miocene 30 Ma (Schlüter, 1997). The mountains stretch from south-east Kenya to south-central Tanzania and are under the direct climatic influence of the Indian Ocean (Fig. 3a). Today, they support 3300–5100 km^2 of tropical forest, which may be less than 30% of the estimated original forested area in prehistoric times (*c.* 2000 years ago) (Newmark, 1998, 2002; Burgess *et al.*, 2007; Platts *et al.*, 2010).

Data

Plant species data were derived from an extensive dataset totalling 75,631 records from the Missouri Botanical Garden's TROPICOS database and from 2216 vegetation plot assessments. Plant species records, representing 3986 vascular plant species, were taxonomically standardized by reference to the African Flowering Plants Database (2008) and further updated by reference to taxonomic revisions and monographs. We also identified all potentially threatened plants on the basis of an assessment by Gereau *et al.* (2010) and all endemic plants on the basis of an analysis by R.E.G. of the plant records from the TROPICOS database.

Funding data were derived from a comprehensive collation of all inventory, research and conservation projects that have taken place in the EAM since 1980 (n = 134), which is when explicit research and conservation interest targeted at the area emerged. All funding data were standardized to US$ in the year 2007 with a GDP deflator (http://www.measuringworth.com). The deflation calculation was made separately for the expenses in every project year. To gain an understanding of the reasons for investment in particular areas, we also compiled criteria for the distribution of funds for all EAM biodiversity survey projects and conservation projects that had a biodiversity survey component since 1980 (n = 61). This was performed on the basis of expert knowledge (N.D.B.) and major donor strategy documents (e.g. CEPF 2003; MNRT Tanzania 2004; WWF-EARPO 2006).

A mechanistic understanding of species richness patterns 'remains the holy grail of modern biogeography and macroecology' (Gotelli *et al.*, 2009), but species richness has been shown to increase with area, available energy (measures include temperature, primary productivity, temperature, potential and actual evapotranspiration), long-term environmental stability and lower levels of disturbance (Whittaker *et al.*, 2001 and references therein). Numbers of threatened and endemic species have been shown to increase with species richness (Jetz *et al.*, 2004), and endemism is also potentially influenced by past climate configurations (Taplin & Lovett, 2003; Jetz *et al.*, 2004; Buckley & Jetz, 2007; Carnaval & Moritz, 2008). Our environmental data therefore included climate, topography, disturbance and distance from the Indian Ocean (a potential proxy for past climatic stability: Hamilton, 1981; Fjeldså *et al.*, 1997; Fjeldså & Lovett, 1997). For all 34 environmental predictors, see Table S1 in Supporting Information. Climatic predictors were derived according to Platts *et al.* (2008), with climate surfaces obtained from the Centre for Resource and Environmental Studies (http://fenner-school.anu.edu.au/). These were then summarized for each mountain bloc. Topography, forest cover data and estimates of the population density were based on Burgess *et al.* (2007). Disturbance per mountain bloc was calculated as the

percentage of trees and poles cut in 949 transects, which totalled 536 km in length and were distributed evenly across mountain blocs. Humans are the single major source of disturbance in the EAM vegetation, as natural disturbances by cyclones, earthquakes, or volcanic eruptions are extremely rare. For the purpose of replication and building on this work, we include a basic data table in the Supporting Information (see Table S2).

Analysis

To establish significant predictors for perceived species richness (*n* species) and numbers of threatened and endemic species, we developed statistical models as follows. In total, 36 candidate predictors were tested: 34 environmental predictors (climate, topography and disturbance); a survey intensity predictor (number of records per mountain bloc); and the cumulative investments per mountain bloc for plant inventories between 1980 and 2007 (see Table S1). Because many of the 24 climatic and the five topographical and forest cover predictors were correlated, principal component analysis (PCA) was used to replace the two variable sets with their uncorrelated components. Hierarchical partitioning (Chevan & Sutherland, 1991) allowed us to estimate the independent and conjoint contributions of all predictors. As a starting point, we fitted a linear regression. Where validation procedures, following Zuur *et al.* (2007), indicated problems associated with heterogeneity of variance, we used linear regression with generalized least squares (GLS) (Pinheiro *et al.*, 2009; Zuur *et al.*, 2009) estimation procedure. GLS was preferred over a Poisson general linear model, as the latter assumes a particular residual distribution that in our case was not matched, and the ranges of our dependent variables were large (i.e. close to continuous). To define the best random structure for the GLS, we first compared models including all starting predictors with different variance covariates (without variance covariate, power of variance covariate, exponential of variance covariate, and constant plus power of a variance covariate) estimated with restricted maximum likelihood, and chose the model with the lowest Akaike Information Criterion (AIC; Sakamoto & Ishiguro, 1986) and the most even spread of residuals. To find the minimum adequate model, we used a backward stepwise selection on the basis of the partial *F*-statistic for regressions, and the likelihood ratio test obtained by maximum likelihood for GLS. Where model validation revealed a Cook's distance greater than one for a single or multiple data points, the analysis was undertaken both with and without extreme observations.

The consistency of the choice of independent variables and the backward stepwise selection procedure model was checked by using two further selection methods where variables had not been replaced by their principal components:

1. Hierarchical partitioning with the full set of predictive variables. Because the hierarchical partitioning function implemented in the R library *hier.part* (Walsh & Mac Nally, 2008) currently only allows for the simultaneous analysis of 12 predictors, we randomly selected 12 predictors for the hierarchical partitioning and averaged the results for each predictor over 100 repetitions. Candidate predictors that had a significantly higher contribution score than the rest of the variables were chosen. Modelling procedures were as above.

2. Stepwise exclusion of predictors based on univariate models. The total set of candidate predictors was reduced to the strongest uncorrelated set (Pearson's $r < 0.7$) according to the predictive power of variables in univariate tests (Quinn & Keough, 2002). This was followed by hierarchical partitioning as above.

The respective contribution of each variable towards explaining the variation in overall perceived plant species richness was established by decomposing the variance in a partial regression (Zuur *et al.*, 2007), whereby for each variable the percentage drop in model fit (R^2) is measured when that variable is omitted from the model. This technique allowed us to establish the contribution of all remaining predictor variables, separately and jointly, towards the level of explained variance. To evaluate the trend over time, the above modelling procedure followed by partitioning of variance was also performed for perceived species richness at decade intervals (1989 and 1999) (with the funding predictor being calculated for the same decade intervals).

Cross-correlations (Chatfield, 2003), used to quantify the association between two variables with a time-lag of *k* years (Zuur *et al.*, 2007), were calculated between the amount of plant inventory funding invested in an area in every year between 1980 and 2007 and the number of new species records for that area, with time-lags ranging from 5 years before to 5 years after. We also calculated cross-correlations between the amount of plant inventory funding and the amount of all other funding (conservation, research, fauna inventories) invested in an area in every year between 1980 and 2007.

We take into consideration that the number of data points used in the analysis was relatively small. However, the dataset contributing to each of these points was extensive. This, in conjunction with the strongly emerging pattern and the consistency established with the model validation procedures, increases our confidence in the reliability of the analysis.

The PCA was calculated in SPSS 11.5; all other statistical analyses were performed in the 'R' statistical and programming environment version 2.9.2 (R Development Core Team 2009) and its libraries *hier.part* (Walsh & Mac Nally, 2008), *nlme* (Pinheiro *et al.*, 2009) and *vegan* (Oksanen *et al.*, 2010).

RESULTS

In the EAM, as in many other parts of the world, funding for biodiversity inventories is scarce. Between 1980 and 2007 investments in conservation and research in the region totalled US$ 117 million (valued at 2007 US$ rate). Of this amount, only 3% has been invested in botanical inventories. Documented vascular plant richness in the EAM, to date totalling 3986 species, has increased in three distinct phases following early explorations pre-1980 (Fig. 1): (1) during the first

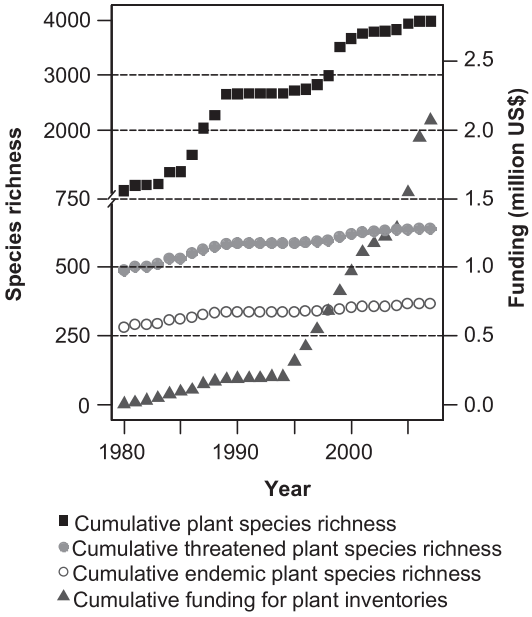

Figure 1 Cumulative perceived species richness and funding in the Eastern Arc Mountains between 1980 and 2007. Funding is for plant inventories and has been standardized to US$ in the year 2007. Note the axis break on the first *Y* axis.

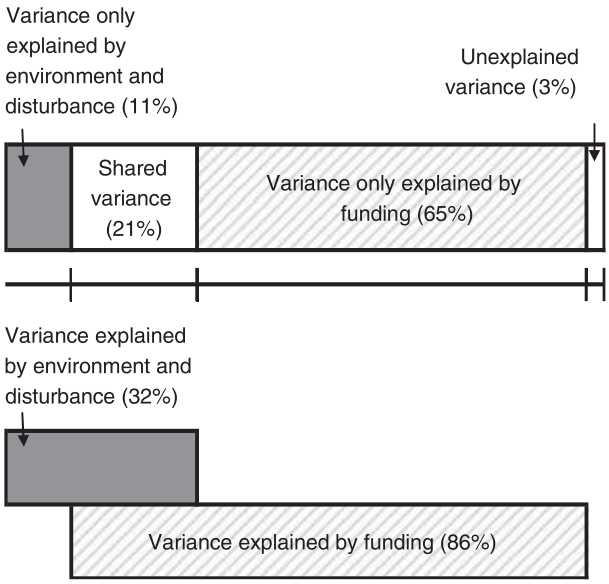

Figure 2 Explained variance in perceived plant species richness for the linear regression model, partitioned between its predictors. Funding is the cumulative investment in plant inventories between 1980 and 2007. The environmental and disturbance predictors are a principal component representative of maximum potential evapotranspiration, evapotranspiration range and maximum suitable temperature days (i.e. water-energy, heterogeneity and optimal growth conditions), and the percentage of trees cut. Independent data points n = 11 (mountain blocs).

intensified exploration in the 1980s; (2) between 1995 and 2000; and (3) from 2004 onwards. The recorded number of threatened and/or endemic species shows a pattern very similar to that of perceived species richness. The funding pattern also shows phases: initial, relatively low levels of funding in the 1980s; a very rapid increase from 1994 to 2001; followed by a second rapid increase from 2004 to present.

Funding (cumulative investments per mountain bloc for plant inventories between 1980 and 2007) emerged as the best predictor for total perceived plant species richness within mountain blocs. This result was consistent across all predictor selection procedures. Partial regressions showed that 65% of the variation in perceived plant species richness is explained by the funding for botanical inventories alone, whereas only 11% is explained by environmental characteristics and disturbance (Fig. 2; see Table S3). The recorded number of threatened and endemic plant species was closely related to overall perceived plant species richness. Both were best predicted by survey intensity in combination with environmental characteristics (see Table S3). Again, the results were consistent across all predictor selection procedures.

Cross-correlations showed distinctive patterns at the bloc resolution level, revealing their different exploration histories. Correlations between funding and perceived species richness with a negative time-lag suggest that perceived species richness is driving funding; conversely, a positive time-lag suggests that funding is driving perceived species richness. A significant correlation for a time-lag of zero also supports the hypothesis that funding is driving perceived species richness levels. This is attributable to the high probability that species will be found relatively quickly once funding has been allocated (for n = 34 projects, we have data on funding start year and year that the majority of the fieldwork took place: for 27 projects, time-lag = 0 year; for seven projects, time-lag = 1 year; average time-lag = 0.21 years ± 0.07 SE), compared to the much lower probability that an increase in perceived richness would trigger the writing of a proposal with funding being allocated for implementation within the same year. The latter is difficult to show as an increase in perceived species richness cannot be assigned to a certain year; however, Fig. S1 shows for example for the Udzungwa Mountains that there was a time-lag of *c.* 5 years between the description of many species from the area (late 1980s) and the allocation of further funding (mid 1990s). For each mountain bloc, cross-correlations were highest at time-lags of zero or greater (Fig. 3b) and indicated overall that funding is a stronger driver of perceived species richness than *vice versa*. Discoveries of threatened and endemic species, which typically require more intensive study, are also likely to be indirectly driven by funding, as both spatial and temporal patterns in perceived threatened, endemic and overall species richness were very similar (see Fig. S2a,b), and survey intensity was partly determined by available funding [Pearson's r (funding, number of records) = 0.7].

The drivers for funding are difficult to disentangle. Our compilation of criteria for the location of investments in biodiversity surveys showed that species-based criteria (perceived species richness and richness in threatened and endemic species) were used in over 80% of the projects. Other criteria

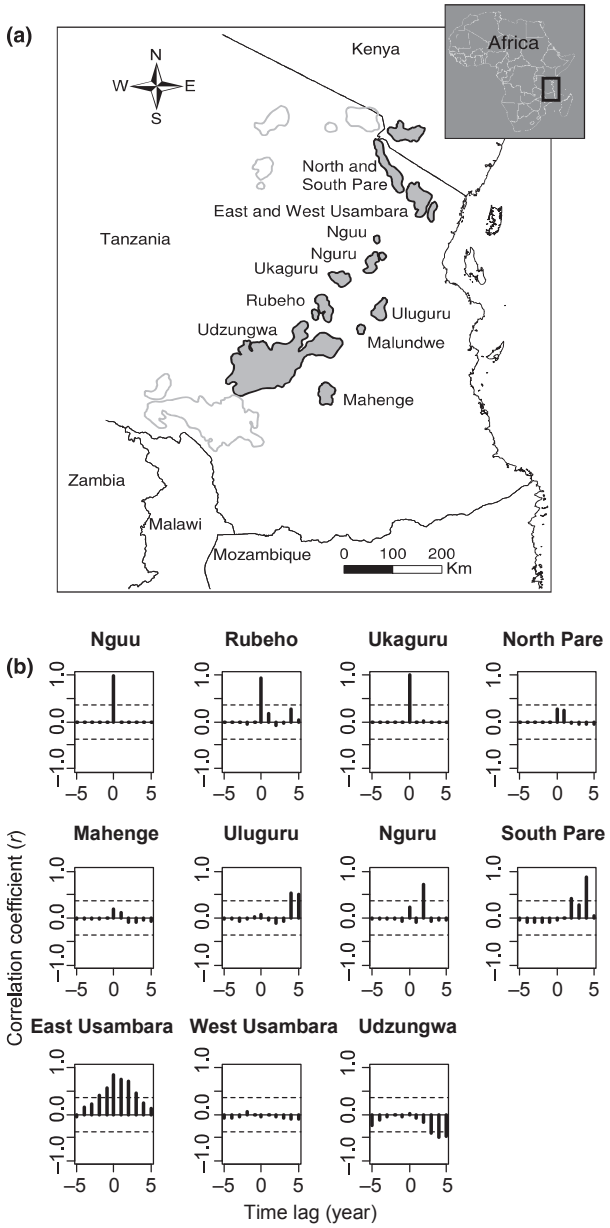

Figure 3 Investments in plant inventories and recorded biodiversity in the Eastern Arc Mountains. (a) Map of the Eastern Arc Mountains. The projection is Universal Transverse Mercator Zone S37, and the datum WGS84. (b) Cross-correlations between investment in plant inventories (1980–2007) and new plant species recorded. Horizontal axis represents time-lag between investment and species discovery, and dotted lines the 95% upper and lower confidence bands. For most mountain blocs, correlations are significant for zero or positive time-lags, suggesting that funding is driving perceived species richness. Correlations are negligible or negative for mountains already well known before the study period where investments did not result in the discovery of further species (West Usambara, Udzungwa) (see Fig. S1 in Supporting Information).

were the level of degradation of the area (30%), the perception of the area as under-researched (25%), and existing infrastructure (5%) (see Fig. S3). Cross-correlations between plant

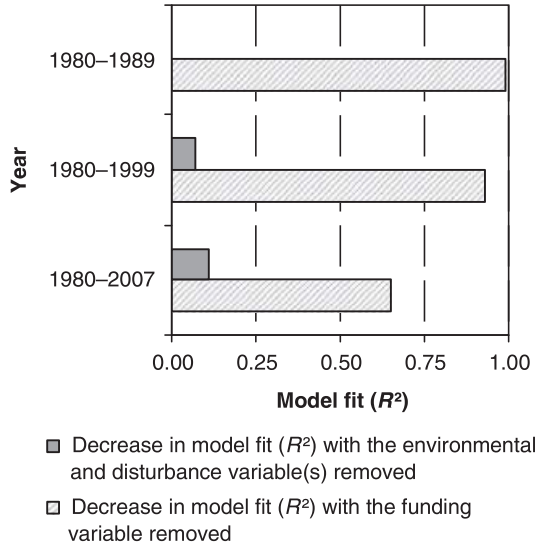

Figure 4 Relative importance of the predictors in explaining perceived plant species richness over time. Funding is the cumulative investment in plant inventories between 1980 and the end of the investigated decade. The environmental and disturbance predictors are a principal component representative of maximum potential evapotranspiration, evapotranspiration range and maximum suitable temperature days and the percentage of trees cut. Independent data points $n = 11$ (mountain blocs).

inventory funding and all other types of funding were varyingly significant for positive, negative or zero time-lags, but the correlations themselves were almost always positive and persisted over a number of years (see Fig. S4). This may indicate that inventory and other funds are at least partly coupled, i.e. that investment triggers further investment for example because of the lower start-up costs of subsequent projects.

The influence of funding on the levels of perceived species richness was strong in the 1980s (Fig. 4), which was the starting period of intensified botanical exploration of the EAM. During this time, a new species for the area would be found with, on average, an investment of less than US$ 100, and funding levels explained a remarkable 99% of variation in the number of new species found (see Table S3). The influence of funding slightly weakened in the 1990s, a phase of highly intense botanical exploration. On average, a new species record was found with every investment of US$ 250, and funding was no longer the sole significant explanatory variable. From 2000 onwards, research in the EAM began to target the lesser researched mountain blocs, resulting in the discovery of 477 new species for the region, 29 of which are potentially threatened and/or endemic. On average, a new species for the area was found with every investment of US$ 500. However, the explanatory power of environmental variables is just above 10%, suggesting that many further botanical inventories will be needed to get a reliable view of species richness patterns in this area. Similarly, patterns in the observed distribution of threatened and endemic species,

which were solely driven by survey intensity in the 1980s, were increasingly related to environmental factors in the 1990s and 2007, with the model coefficient for the survey intensity predictor steadily decreasing (see Table S3). Funds for conservation, research and inventory were highly correlated with each other (see Table S4), showing that the same mountain blocs are targeted for these different purposes.

DISCUSSION

This study implies that we may have a much distorted view of species diversity patterns and may not have sufficient data to identify conservation priority areas with certainty, particularly at the site scale (Da Fonseca *et al.*, 2000). In the EAM, as elsewhere, research and conservation investment is biased towards the areas that we think are important, and our understanding of the relative conservation importance of areas within different biomes across the world may be biased towards those that have received the most funding for biodiversity inventories. This may also partially explain why surprisingly little consensus has been achieved on the distributional pattern and drivers of species richness (Rahbek *et al.*, 2007; Gotelli *et al.*, 2009), and possibly also patterns of endemism, which have, for example, been attributed to evolutionary effects and historical climate configurations (Taplin & Lovett, 2003; Jetz *et al.*, 2004; Buckley & Jetz, 2007; Carnaval & Moritz, 2008).

The close tie between the funds for inventory and those for conservation suggests that where initial biodiversity inventories result in the discovery of new species, further funding is then attracted for conservation and research. This partly finances further inventories and results in the discovery of more species for that area, strengthening its value for conservation and research in a spiral fashion. In their study of global patterns of plant diversity, Kier *et al.* (2005), for example, have shown that species-rich ecoregions are better inventoried than areas poor in vascular plants. Similarly, in their re-definition of African hotspots Küper *et al.* (2004) suggest that areas perceived as species-rich may have attracted further biodiversity surveys with their conservation status thereby amplified. Reddy & Davalos (2003) find that sampling of birds in sub-Saharan Africa has been significantly concentrated within and around areas now designated as conservation priorities. At the same time, other areas receive relatively little initial funding for inventories; hence, the number of species discovered in these areas remains small and their conservation and research status low.

The reasons for the investment of further funding in already relatively well-researched areas are difficult to disentangle. Projects may apply conservation criteria such as levels of perceived biodiversity, but they may also continue to focus on the same few areas today because an existing infra-structure of roads, field stations and institutional contacts greatly reduces the cost of future projects. Furthermore, donors may continue to invest in the same area for historical reasons or simply to avoid overlap with other donors. Because of the many inter-correlations between these factors, it is difficult to deduce causal relationships with any certainty, but major biodiversity conservation donor strategy documents for the region (e.g. CEPF (Critical Ecosystem Partnership Fund), 2003; MNRT Tanzania (Ministry of Natural Resources and Tourism Tanzania), 2004; WWF-EARPO, 2006) show that perceived biodiversity criteria (species richness and numbers of endemic and threatened species), in combination with other considerations such as cost-effectiveness (existing infrastructure and/or research niches, i.e. under-researched areas) and the area's level of threat, are the major factors applied in funding prioritization decisions.

The selection of initial areas for investment is often related to accessibility, interest in the area, size, (historical/colonial) land ownership and political considerations (Reddy & Davalos, 2003; Kadmon *et al.*, 2004; Halpern *et al.*, 2006). The positive association between perceived species richness and the percentage of trees cut in the EAM, for example, may be a reflection of nonlinear effects of disturbance on species richness, but it may also be because of easily accessible areas close to roads, and markets are targeted by both logging companies and botanists, as historically remote areas would have been extremely difficult if not impossible to reach. This may mean that both investments and perceived levels of biodiversity are biased towards areas that are easier to access.

Our study is regional in scope, which is the scale at which many conservation decisions take place (Mace *et al.*, 2000; Ferrier, 2002; Ferrier *et al.*, 2004). Further study in other areas is needed to establish whether the funding–biodiversity circularity we find here holds true elsewhere. Implications for global conservation prioritization schemes may be limited, as these are typically based on expert opinion (Brooks *et al.*, 2006), and it is difficult to imagine that priority biodiversity areas could have been overlooked at this scale. However, factors such as political instability mean that global survey intensity is unequally distributed, with, for example, regions within Afghanistan, Angola, Colombia, Iraq, Somalia, Sudan, the Central African Republic and the Democratic Republic of Congo being poorly collected as a result of extremely challenging research conditions. The conservation value of many of these areas is insufficiently documented or unknown (Küper *et al.*, 2004, 2006, Dr. Matthew Hall, Royal Botanic Garden Edinburgh, personal communication).

Promising progress has been made with species distribution modelling techniques (Elith *et al.*, 2006), which can be used to establish the probability of an area's conservation importance for one or several species (Da Fonseca *et al.*, 2000; Graham *et al.*, 2004) or to guide future biodiversity surveys e.g. towards areas with the largest difference between recorded biodiversity and expected biodiversity (Küper *et al.*, 2006). Climatically based fine-scale species distribution models for the area (Platts *et al.*, 2010), in line with individual-based rarefaction curves (see Fig. S5) for relatively well-sampled mountain blocs with >3000 records ($n = 6$), suggest that for example the conservation importance of the Rubeho and Nguru Mountains may be underestimated. However, the predictions from these

models for new areas and future scenarios will only be as good as the data that underpin them (Rondinini *et al.*, 2006). We recommend that funding and associated sampling intensity biases be considered in the development of these models to achieve more accurate predictions, or at least to inform relevant measures of uncertainty. Promise is also held by the various diversity indices, species richness estimators (extrapolation from species accumulation curves, parametric methods and nonparametric estimators) and rarefaction curves that allow for the comparison of diversity levels of differently intensively sampled sites (Magurran, 2004). However, all of these techniques are sensitive to sampling effort below a certain minimum sample size (Magurran, 2004). Gimaret-Carpentier *et al.* (1998) and Lande *et al.* (2000), for example, find the Simpson indicator to be very robust to low sample sizes and recommend a minimum sample size of 300–400 individuals per investigated site in moist evergreen forest; Sorensen *et al.* (2002) estimate that the needed sample size for reasonable estimates of species richness in high diversity areas is 30–50:1 (individuals:number of species). A quarter of the EAM blocs have <300 records (and over half < 5000), i.e. it would be inappropriate to attempt to estimate actual total species richness for these. These methods are further confounded by differences in the detectability of species, in observer efficiency, in the underlying species abundance distribution and/or in habitat heterogeneity as e.g. caused by disturbance (Lande *et al.*, 2000; Magurran, 2004).

In conclusion, our study implies that there may be a need for a more balanced distribution of conservation and inventory investments as lesser known areas may be underestimated in their conservation importance. We are not advocating a highly cost-intensive global standardization of survey intensity for all areas and also recognize that the start-up costs for surveys in hitherto under-researched areas can be high. We also agree that in the face of rapid global biodiversity loss, it is important to dedicate efforts to conservation even if our knowledge of patterns in species richness is incomplete (Meir *et al.*, 2004; McDonald-Madden *et al.*, 2008; Grantham *et al.*, 2009) and that species-based conservation approaches are severely limited in their scope anyway, with stronger emphasis needed on preserving processes that generate and sustain biodiversity (Cowling *et al.*, 2004; Pressey, 2004; Knight *et al.*, 2007). Furthermore, we must ask whether investments into the collection of biodiversity data will indeed increase the effectiveness of conservation planning; Grantham *et al.* (2008) have shown that further investments in data collection can have rapidly diminishing returns in terms of improved conservation planning; Cowling *et al.* (2010) have suggested that for complex and species-rich systems, conservation funds may be more suitably directed towards improved management of already gazetted areas (New, 2006), or towards the mapping of socio-economic data and conservation costs, i.e. restrictions to conservation opportunity, which can exhibit greater spatial heterogeneity than biodiversity itself (Knight & Cowling, 2007; Bode *et al.*, 2008; Guerrero *et al.*, 2010; Knight *et al.*, 2010). However, we take the view that further studies across several spatial scales and based on complete biodiversity inventories for several taxonomic groups will be needed to establish whether, overall, conservation planning based on biodiversity inventory data is or is not more effective than conservation planning based on opportunity alone. By dedicating funds nearly exclusively towards areas that are already perceived as important and/or where project infra-structure already exists, we may risk losing other areas of equal importance with greatly underestimated or unknown conservation status (Kareiva & Marvier, 2003; Cowling *et al.*, 2010). Increasing the funding for biodiversity inventories in potentially important but under-researched areas, we think, would reduce this risk and provide a more balanced assessment of diversity patterns, allowing effective conservation of more of the world's biodiversity.

ACKNOWLEDGEMENTS

We thank A. Balmford, S.A. Bhagwat, M. Hall, P.M. Hollingsworth, A.T. Knight, R.T. Pennington, D. Raffaelli and one anonymous reviewer for critical discussions. Funding for A.A. and P.J.P. was provided by a Marie Curie Actions grant to R.M. (MEXT-CT-2004-517098), and for M.T.B. by NERC (NE/603 E006795/1). The authors gratefully acknowledge contributions of plant, disturbance and funding data: Frontier Tanzania, the Tanzania Forest Conservation Group, the Tanzanian Forestry and Beekeeping Division, A.R. Marshall, and J.B. Hall provided plant and disturbance field data. BirdLife International (G. Eshiamwata, P. Ndang'ang'a, and P. Nnyiti,), the Critical Ecosystem Partnership Fund (G. Love), DANIDA (T. Blomley), T. Evans, FINNIDA (M. Määttä, M. Nummelin, J. Saramäki, and V. Turpeien), GTZ (Steffen Rötcher), G. Hertel, Missouri Botanical Garden (Peter Phillipson and Porter P. Lowry II), Museo Tridentino di Scienze Naturali (M. Menegon, F. Rovero), B. Newmark, T. Struhsaker, The Field Museum of Natural History Chicago (B. Stanley), TFCG (N. Doggart), the UNDP/GEF Cross Border Biodiversity Programme (L. Persha), University of Copenhagen (J. Fjeldså and L. Hansen), University of York (A.R. Marshall) and WWF (P.J. Stephenson and P. Sumbi) provided funding data.

REFERENCES

African Flowering Plants Database (2008) Conservatoire et Jardin botaniques de la Villle de Genève and South African National Biodiversity Institute Pretoria. Available at: http://www.ville-ge.ch/musinfo/bd/cjb/africa/ (accessed October 2008).

Bode, M., Wilson, K.A., Brooks, T.M., Turner, W.R., Mittermeier, R.A., McBride, M.F., Underwood, E.C. & Possingham, H.P. (2008) Cost-effective global conservation spending is robust to taxonomic group. *Proceedings of the National Academy of Sciences USA*, **105**, 6498–6501.

Brooks, T.M., Mittermeier, R.A., da Fonseca, G.A.B., Gerlach, J., Hoffmann, M., Lamoreux, J.F., Mittermeier, C.G., Pilgrim,

J.D. & Rodrigues, A.S.L. (2006) Global biodiversity conservation priorities. *Science*, **313**, 58–61.

Buckley, L.B. & Jetz, W. (2007) Environmental and historical constraints on global patterns of amphibian richness. *Proceedings of the Royal Society B: Biological Sciences*, **274**, 1167–1173.

Burgess, N.D., Butynski, T.M., Cordeiro, N.J., Doggart, N.H., Fjelds, J.å., Howell, K.M., Kilahama, F.B., Loader, S.P., Lovett, J.C., Mbilinyi, B., Menegon, M., Moyer, D.C., Nashanda, E., Perkin, A., Rovero, F., Stanley, W.T. & Stuart, S.N. (2007) The biological importance of the Eastern Arc Mountains of Tanzania and Kenya. *Biological Conservation*, **134**, 209–231.

Carnaval, A.C. & Moritz, C. (2008) Historical climate modelling predicts patterns of current biodiversity in the Brazilian Atlantic forest. *Journal of Biogeography*, **35**, 1187–1201.

CEPF (Critical Ecosystem Partnership Fund) (2003). *Eastern Arc Mountains and coastal forest hotspot Ecosystem Profile*. CEPF, Dar es Salaam, Nairobi.

Chatfield, C.. (2003) *The analysis of time series: an introduction*. Chapman and Hall, London.

Chevan, A. & Sutherland, M. (1991) Hierarchical partitioning. *American Statistician*, **45**, 90–96.

Cowling, R.M., Knight, A.T., Faith, D.P., Ferrier, S., Lombard, A.T., Driver, A., Rouget, M., Maze, K. & Desmet, P.G. (2004) Nature conservation requires more than a passion for species. *Conservation Biology*, **18**, 1674–1676.

Cowling, R.M., Knight, A.T., Privett, S.D.J. & Sharma, G.C.B. (2010) Invest in opportunity, not inventory of hotspots. *Conservation Biology*, **24**, 633–635.

Da Fonseca, G.A.B., Balmford, A., Bibby, C., Boitani, L., Corsi, F., Brooks, T., Gascon, C., Olivieri, S., Mittermeier, R.A., Burgess, N., Dinerstein, E., Olson, D., Hannah, L., Lovett, J., Moyer, D., Rahbek, C., Stuart, S. & Williams, P. (2000) It's time to work together and stop duplicating conservation efforts... following Africa's lead in setting priorities. *Nature*, **405**, 393–394.

Dalton, R. (2000) Ecologists back blueprint to save biodiversity hotspots. *Nature*, **406**, 926.

Eken, G., Bennun, L., Brooks, T.M., Darwall, W., Fishpool, L.D.C., Foster, M., Knox, D., Langhammer, P., Matiku, P., Radford, E., Salaman, P., Sechrest, W., Smith, M.L., Spector, S. & Tordoff, A. (2004) Key biodiversity areas as site conservation targets. *BioScience*, **54**, 1110–1118.

Elith, J., Graham, C.H., Anderson, R.P. *et al.* (2006) Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, **29**, 129–151.

Ferrier, S. (2002) Mapping spatial pattern in biodiversity for regional conservation planning: where to from here? *Systematic Biology*, **51**, 331–363.

Ferrier, S., Powell, G.V.N., Richardson, K.S., Manion, G., Overton, J.M., Allnutt, T.F., Cameron, S.E., Mantle, K., Burgess, N.D., Faith, D.P., Lamoreux, J.F., Kier, G., Hijmans, R.J., Funk, V.A., Cassis, G.A., Fisher, B.L., Flemons, P., Lees, D., Lovett, J.C. & Van Rompaey, R. (2004) Mapping more of terrestrial biodiversity for global conservation assessment. *BioScience*, **54**, 1101–1109.

Fishpool, L.D.C. & Evans, M.I. (2001) *Important bird areas in Africa and associated islands*. BirdLife International, Cambridge.

Fjeldså, J. & Lovett, J.C. (1997) Biodiversity and environmental stability. *Biodiversity and Conservation*, **6**, 315–323.

Fjeldså, J., Ehrlich, D., Lambin, E. & Prins, E. (1997) Are biodiversity 'hotspots' correlated with current ecoclimatic stability? A pilot study using the NOAA-AVHRR remote sensing data. *Biodiversity and Conservation*, **6**, 401–422.

Gereau, R.E., Taylor, C.M., Bodine, S. & Kindeketa, W.J. (2010) Plant conservation assessment in the Eastern Arc Mountains and coastal forests of Tanzania and Kenya. http://www.mobot.org/MOBOT/Research/tanzania/cepf.shtml.

Gimaret-Carpentier, C., Pelissier, R., Pascal, J.P. & Houllier, F. (1998) Sampling strategies for the assessment of tree species diversity. *Journal of Vegetation Science*, **9**, 161–172.

Gotelli, N.J., Anderson, M.J., Arita, H.T. *et al.* (2009) Patterns and causes of species richness: a general simulation model for macroecology. *Ecology Letters*, **12**, 873–886.

Graham, C.H., Ferrier, S., Huettman, F., Moritz, C. & Peterson, A.T. (2004) New developments in museum-based informatics and applications in biodiversity analysis. *Trends in Ecology and Evolution*, **19**, 497–503.

Grantham, H.S., Moilanen, A., Wilson, K.A., Pressey, R.L., Rebelo, T.G. & Possingham, H.P. (2008) Diminishing return on investment for biodiversity data in conservation planning. *Conservation Letters*, **1**, 190–198.

Grantham, H.S., Wilson, K.A., Moilanen, A., Rebelo, T. & Possingham, H.P. (2009) Delaying conservation actions for improved knowledge: how long should we wait? *Ecology Letters*, **12**, 293–301.

Guerrero, A.M., Knight, A.T., Grantham, H.S., Cowling, R.M. & Wilson, K.A. (2010) Predicting willingness-to-sell and its utility for assessing conservation opportunity for expanding protected area networks. *Conservation Letters*, **3**, 332–339.

Halpern, B.S., Pyke, C.R., Fox, H.E., Haney, J.C., Schlaepfer, M.A. & Zaradic, P. (2006) Gaps and mismatches between global conservation priorities and spending. *Conservation Biology*, **20**, 56–64.

Hamilton, A.C. (1981) The quaternary history of African forests – its relevance to conservation. *African Journal of Ecology*, **19**, 1–6.

Hawkins, B.A., Field, R., Cornell, H.V., Currie, 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.

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.

Kadmon, R., Farber, O. & Danin, A. (2004) Effect of roadside bias on the accuracy of predictive maps produced by bioclimatic models. *Ecological Applications*, **14**, 401–413.

Kareiva, P. & Marvier, M. (2003) Conserving biodiversity coldspots – Recent calls to direct conservation funding to the world's biodiversity hotspots may be bad investment advice. *American Scientist*, **91**, 344–351.

Kier, G., Mutke, J., Dinerstein, E., Ricketts, T.H., Kuper, W., Kreft, H. & Barthlott, W. (2005) Global patterns of plant diversity and floristic knowledge. *Journal of Biogeography*, **32**, 1107–1116.

Knight, A.T. & Cowling, R.M. (2007) Embracing opportunism in the selection of priority conservation areas. *Conservation Biology*, **21**, 1124–1126.

Knight, A.T., Smith, R.J., Cowling, R.M., Desmet, P.G., Faith, D.P., Ferrier, S., Gelderblom, C.M., Grantham, H., Lombard, A.T., Maze, K., Nel, J.L., Parrish, J.D., Pence, G.Q.K., Possingham, H.P., Reyers, B., Rouget, M., Roux, D. & Wilson, K.A. (2007) Improving the key biodiversity areas approach for effective conservation planning. *BioScience*, **57**, 256–261.

Knight, A.T., Cowling, R.M., Difford, M. & Campbell, B.M. (2010) Mapping human and social dimensions of conservation opportunity for the scheduling of conservation action on private land. *Conservation Biology*, **24**, 1348–1358.

Küper, W., Sommer, J.H., Lovett, J.C., Mutke, J., Linder, H.P., Beentje, H.J., Van Rompaey, R., Chatelain, C., Sosef, M. & Barthlott, W. (2004) Africa's hotspots of biodiversity redefined. *Annals of the Missouri Botanical Garden*, **91**, 525–535.

Küper, W., Sommer, J.H., Lovett, J.C. & Barthlott, W. (2006) Deficiency in African plant distribution data - missing pieces of the puzzle. *Botanical Journal of the Linnean Society*, **150**, 355–368.

Lande, R., DeVries, P.J. & Walla, T.R. (2000) When species accumulation curves intersect: implications for ranking diversity using small samples. *Oikos*, **89**, 601–605.

Mace, G.M., Balmford, A., Boitani, L., Cowlishaw, G., Dobson, A.P., Faith, D.P., Gaston, K.J., Humphries, C.J., Vane-Wright, R.I., Williams, P.H., Lawton, J.H., Margules, C.R., May, R.M., Nicholls, A.O., Possingham, H.P., Rahbek, C. & van Jaarsveld, A.S. (2000) It's time to work together and stop duplicating conservation efforts. *Nature*, **405**, 393.

Magurran, A.E. (2004) *Measuring biological diversity*. Blackwell, Oxford.

McDonald-Madden, E., Bode, M., Game, E.T., Grantham, H. & Possingham, H.P. (2008) The need for speed: informed land acquisitions for conservation in a dynamic property market. *Ecology Letters*, **11**, 1169–1177.

Meir, E., Andelman, S. & Possingham, H.P. (2004) Does conservation planning matter in a dynamic and uncertain world? *Ecology Letters*, **7**, 615–622.

Mittermeier, R.A., Robles-Gil, P., Hoffmann, M., Pilgrim, J.D., Brooks, T.M., Mittermeier, C.G., Lamoreux, J.L. & Fonseca, G. (2005) *Hotspots revisited: Earth's biologically richest and most endangered ecoregions*, Cemex, Mexico City.

MNRT Tanzania (Ministry of Natural Resources and Tourism Tanzania) (2004). *Biodiversity conservation thematic strategy*. Eastern Arc Mountains Strategy. Conservation and Management of the Eastern Arc Mountain Forests of, Tanzania. GEF-UNDP-URT/01/00015426. Available at: http://easternarc.or.tz/downloads/Thematic-strategies/Biodiversity%20 strategy.PDF (accessed July 2010).

Nelson, B.W., Ferreira, C.A.C., Dasilva, M.F. & Kawasaki, M.L. (1990) Endemism centers, refugia and botanical collection density in Brazilian Amazonia. *Nature*, **345**, 714–716.

New, T.R. (2006) Does inventory of insect species in the tropics really help in their conservation? *Journal of Insect Conservation*, **10**, 211–213.

Newmark, W. (1998) Forest area, fragmentation and loss in the Eastern Arc Mountains: implications for the conservation of biological diversity. *Journal of East African Natural History*, **87**, 29–36.

Newmark, W.D. (2002) *Conserving biodiversity in East African forests: A study of the Eastern Arc Mountains*. Springer-Verlag, Berlin.

O'Brien, E.M., Field, R. & Whittaker, R.J. (2000) Climatic gradients in woody plant (tree and shrub) diversity: water-energy dynamics, residual variation, and topography. *Oikos*, **89**, 588–600.

Oksanen, J., Blanchet, F.G., Kindt, R., Legendre, P., O'Hara, R.B., Simpson, G.L., Solymos, P., Stevens, M.H.H. & Wagner, H. (2010) *vegan: Community Ecology Package*. R package version 1.17-2. R Development Core Team, Vienna. Available at: http://CRAN.R-project.org/package=vegan (accessed June 2010).

Olson, D.M. & Dinerstein, E. (2002) The Global 200: priority ecoregions for global conservation. *Annals of the Missouri Botanical Garden*, **89**, 199–224.

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D. & the R Core Team (2009) nlme: linear and nonlinear mixed effects models. R package version 3.1-96. R Development Core Team, Vienna. Available at: http://www.stats.bris.ac.uk/R/ (accessed January 2009).

Plantlife International (2004) *Identifying and protecting the world's most important plant areas*. Plantlife International, Salisbury.

Platts, P.J., McClean, C.J., Lovett, J.C. & Marchant, R. (2008) Predicting tree distributions in an East African biodiversity hotspot: model selection, data bias and envelope uncertainty. *Ecological Modelling*, **218**, 121–134.

Platts, P.J., Ahrends, A., Gereau, R.E., McClean, C., Lovett, J.C., Marshall, A., Pellikka, P., Mulligan, M., Fanning, E. & Marchant, R. (2010) Can distribution models help refine inventory-based estimates of conservation priority? A case study in the Eastern Arc forests of Tanzania and Kenya. *Diversity and Distributions*, **16**, 628–642.

Pressey, R.L. (2004) Conservation planning and biodiversity: assembling the best data for the job. *Conservation Biology*, **18**, 1677–1681.

Quinn, Q.P. & Keough, M.J. (2002) *Experimental design and data analysis for biologists*. Cambridge University Press, Cambridge.

R Development Core Team (2009) *R: a language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna. Available at: http://www.R-project.org (accessed June 2009).

Rahbek, C., Gotelli, N.J., Colwell, R.K., Entsminger, G.L., Rangel, T. & 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.

Reddy, S. & Davalos, L.M. (2003) Geographical sampling bias and its implications for conservation priorities in Africa. *Journal of Biogeography*, **30**, 1719–1727.

Rondinini, C., Wilson, K.A., Boitani, L., Grantham, H. & Possingham, H.P. (2006) Tradeoffs of different types of species occurrence data for use in systematic conservation planning. *Ecology Letters*, **9**, 1136–1145.

Sakamoto, Y. & Ishiguro, M.K.G. (1986) *Akaike Information Criterion Statistics, D*. Reidel Publishing Company, Dordrecht.

Schlüter, T. (1997) *Geology of East Africa*, Borntraeger, Berlin.

Sorensen, L.L., Coddington, J.A. & Scharff, N. (2002) Inventorying and estimating subcanopy spider diversity using semiquantitative sampling methods in an Afromontane forest. *Environmental Entomology*, **31**, 319–330.

Soria-Auza, R.W. & Kessler, M. (2008) The influence of sampling intensity on the perception of the spatial distribution of tropical diversity and endemism: a case study of ferns from Bolivia. *Diversity and Distributions*, **14**, 123–130.

Stattersfield, A.J., Crosby, M.J., Long, A.J. & Wege, D.C. (1998) *Endemic bird areas of the world: Priorities for biodiversity conservation*, BirdLife International, Cambridge.

Taplin, J.R.D. & Lovett, J.C. (2003) Can we predict centres of plant species richness and rarity from environmental variables in sub-Saharan Africa? *Botanical Journal of the Linnean Society*, **142**, 187–197.

Walsh, C. & Mac Nally, R. (2008). hier.part: Hierarchical Partitioning. R package version 1.0-3. Available at: http://www.stats.bris.ac.uk/R/ (accessed January 2009).

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

WWF-EARPO (Worldwide Fund for Nature Eastern African Regional Programme Office) (2006) *Coastal Forests of Eastern Africa: Conservation strategy and action plan*. WWF Eastern Africa Regional Programme Office, Nairobi.

Zuur, A.F., Ieno, E.N. & Smith, G.M. (2007) *Analysing ecological data*, Springer, New York.

Zuur, A.F., Ieno, E.N., Walker, N.J., Saveliev, A.A. & Smith, G.M. (2009) *Mixed effects models and extension in ecology with R*, Springer, New York.

SUPPORTING INFORMATION

Additional Supporting information may be found in the online version of this article:

Figure S1 Levels of funding, number of specimens and discoveries of new and threatened species over time in individual mountain blocs.

Figure S2 Cross-correlations between investment in plant inventories (1980–2007) and new threatened and endemic plant species recorded.

Figure S3 Criteria used for the allocation of funding across Eastern Arc Mountain blocs for biodiversity survey projects.

Figure S4 Cross-correlations between investment in plant inventories and all other investments (1980–2007).

Figure S5 Individual-based rarefaction curves for relatively well-sampled mountain blocs with > 3000 records.

Table S1 Starting environmental and disturbance predictors used in the modelling process.

Table S2 Raw data table.

Table S3 Minimum adequate models for perceived plant species richness, and levels of threatened and endemic plant species richness.

Table S4 Correlations between total funding (1980–2007) for inventories, conservation and other research in the Eastern Arc Mountains.

As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials are peer-reviewed and may be re-organized for online delivery, but are not copy-edited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.

BIOSKETCH

Antje Ahrends is a plant conservation scientist based at the Royal Botanic Garden of Edinburgh. Her work within the Genetics and Conservation Programme (http://www.rbge.org.uk/science/genetics-and-conservation) focuses on the impact of environmental change on plant diversity and forest condition, and the influence of data quality and biases on conservation priority area selection. She was previously based at the York Institute for Tropical Ecosystem Dynamics (KITE) – a Marie-Curie funded Excellence Centre that explores the relationship between vegetation dynamics, climate change and human impacts in Africa (http://www.york.ac.uk/res/kite/).

Author contributions: C.R. and A.A. conceptualized the study. A.A. and M.T.B. conducted all the analyses, on the basis of data provided by N.D.B., R.E.G., P.J.P., R.M., J.C.L., V.W.K., N.O. and E.F. A.A. wrote the paper, and all authors discussed the results and commented on the manuscript. A.A. and P.J.P. were funded by a grant to R.M.

Editor: Richard Cowling