

# Global patterns and determinants of vascular plant diversity

Holger Kreft\*<sup>†</sup> and Walter Jetz\*

\*Nees Institute for Biodiversity of Plants, University of Bonn, Meckenheimer Allee 170, D-53115 Bonn, Germany; and <sup>†</sup>Division of Biological Sciences, University of California at San Diego, 9500 Gilman Drive MC 0116, La Jolla, CA 92093-0116

Edited by F. Stuart Chapin III, University of Alaska, Fairbanks, AK, and approved January 25, 2007 (received for review September 22, 2006)

**Plants, with an estimated 300,000 species, provide crucial primary production and ecosystem structure. To date, our quantitative understanding of diversity gradients of megadiverse clades such as plants has been hampered by the paucity of distribution data. Here, we investigate the global-scale species-richness pattern of vascular plants and examine its environmental and potential historical determinants. Across 1,032 geographic regions worldwide, potential evapotranspiration, the number of wet days per year, and measurements of topographical and habitat heterogeneity emerge as core predictors of species richness. After accounting for environmental effects, the residual differences across the major floristic kingdoms are minor, with the exception of the uniquely diverse Cape Region, highlighting the important role of historical contingencies. Notably, the South African Cape region contains more than twice as many species as expected by the global environmental model, confirming its uniquely evolved flora. A combined multipredictor model explains  $\approx 70\%$  of the global variation in species richness and fully accounts for the enigmatic latitudinal gradient in species richness. The models illustrate the geographic interplay of different environmental predictors of species richness. Our findings highlight that different hypotheses about the causes of diversity gradients are not mutually exclusive, but likely act synergistically with water–energy dynamics playing a dominant role. The presented geostatistical approach is likely to prove instrumental for identifying richness patterns of the many other taxa without single-species distribution data that still escape our understanding.**

biodiversity | historical contingency | latitudinal gradient | macroecology | species richness

Geographic patterns of species distributions are central to ecology (1–7). Progress toward more general and, importantly, global models of gradients of species richness to date has been hampered by the many species that remain only poorly documented in their geographic occurrence or altogether unknown (8). For an understanding of the global distribution of diversity, plants might be of particular relevance. Plants comprise some 300,000 species, are key structural elements of terrestrial ecosystems, and are the basis of all terrestrial food webs. High plant diversity is likely associated with high biotic heterogeneity and thus a higher potential for specialization in various animal groups (9). Generally, medium to strong positive relationships between producer and consumer diversity have been found (10–12). Plants may thus play a central role as an indicator group; under this assumption, their richness pattern has already been used extensively for global-scale conservation priority setting (13).

Recently, considerable progress has been made toward documenting broad-scale patterns of plant richness (11, 14–21). In general, two different data-type approaches are possible to map and analyze global richness gradients (22). First, studies may be based on single-species occurrence data in the form of locality records or expert range maps (23–25). Unfortunately, this approach is limited by the relatively small fraction of all species for which such data are available. Second, and consequently, a

single-species approach for mapping and analyzing global distributions of many speciose groups, such as vascular plants, will long remain elusive. Thus, the method of choice is analyzing the species-richness information for geographic units with floras that have been well described (11, 21, 22). This method offers a powerful approach to understanding the variability of plant species richness at a global scale.

A number of studies have shown a remarkably strong association between contemporary climate and species richness (4, 5, 7, 11, 16, 26–28). According to the “water–energy dynamics hypothesis,” species richness at higher latitudes is controlled by the availability of ambient heat, whereas, in the thermally suitable tropics, water- and humidity-related variables are the main driving factors (4, 5, 28). Alternatively or additionally, the sensitivity of most plants to frost or drought may constrain their richness outside warm and humid regions (5, 29). Another set of hypotheses states that habitat heterogeneity governs species-richness gradients by local and regional species turnover (30, 32). Third, historical/evolutionary hypotheses attribute species-richness gradients to geographic differences in the geological and climatic history, such as tectonic movements, uplift of mountain ranges, long-term climatic stability, or Pleistocene cooling and dryness. These historic factors may cause divergent rates of diversification (32–35). Recently, water- and energy-related variables have been found to be dominant predictors of global angiosperm family richness (27). Whether this would question a major role for historical factors has been debated (35). To date, relative roles of potential environmental and historical drivers of species diversity lack scrutiny.

Here, we present an analysis of geographic patterns and putative macroecological determinants of vascular plant diversity at the species level and with a global scope. Our analysis is based on an exhaustive data set of 1,032 regions worldwide that has been used to produce expert opinion-based continental to global maps of plant species richness (11, 14, 17, 21). We use both nonspatial and spatial (controlling for spatial autocorrelation) modeling techniques to test, in turn, the predictive potential of variables representing different hypotheses. We proceed to develop a combined multipredictor model and use it in conjunction with geostatistical techniques to predict vascular plant diversity across the whole world. We thereby outline a general geostatistical approach to capture the richness gradients of the many less studied groups of organisms that still escape our understanding.

Author contributions: H.K. and W.J. designed research; H.K. and W.J. performed research; H.K. and W.J. analyzed data; and H.K. and W.J. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Abbreviations: AIC, Akaike information criterion; GLM, generalized linear model; PET, potential evapotranspiration; SLM, spatial linear model.

<sup>†</sup>To whom correspondence should be addressed. E-mail: hk@uni-bonn.de.

This article contains supporting information online at [www.pnas.org/cgi/content/full/0608361104/DC1](http://www.pnas.org/cgi/content/full/0608361104/DC1).

© 2007 by The National Academy of Sciences of the USA

**Table 1. Results of GLM and SLM for selected predictor variables and species richness of vascular plants**

Hypothesis and model	GLM				SLM			
	<i>t</i>	Deviance, %	AIC	Moran's <i>I</i>	<i>z</i>	Deviance, %	AIC	Moran's <i>I</i>
NULL	—	—	738	0.43	—	39.2	306	—
AREA	8.5	6.6	669	0.53	20.2***	57.4	−24	−0.02
Energy								
PET	14.2	25.4 (16.3)	439 (556)	0.34	6.9***	58.7 (40.6)	−72 (265)	−0.02
TEMP	9.8	16.9 (8.5)	551 (648)	0.37	13.6*	57.3 (39.1)	−22 (303)	−0.02
Water								
PRE	17.6	35.9 (23.2)	283 (467)	0.33	11.1***	63.9 (44.1)	−216 (200)	−0.03
WETDAYS	8.9	16.5 (7.2)	556 (663)	0.43	9.1***	64.2 (43.8)	−196 (228)	−0.03
Water–energy								
AET	20.3	41.4 (28.6)	191 (391)	0.28	13.1***	63.8 (44.2)	−224 (187)	−0.02
WAT-ENER	20.2	41.3 (28.4)	193 (394)	0.29	12.8***	64.5 (45.4)	−241 (169)	−0.02
PET + WETDAYS		51.1 (34.9)	6 (299)	0.24		65.8 (46.1)	−296 (145)	−0.02
PET	20.9	34.6 (27.7)			11.5***			
WETDAYS	17.1	25.7 (18.5)			11.9***			
PET × WETDAYS		52.6 (36.8)	−24 (270)	0.21		65.8 (46.1)	−297 (143)	−0.02
PET	−3.9	34.6 (27.7)			−0.8			
WETDAYS	−4.6	25.7 (18.5)			−1.3			
PET:WETDAYS	5.6	1.5 (1.9)			1.99*			
Heterogeneity								
TOPOVEG	20.3	28.7 (28.6)	392 (392)	0.50	24.4***	63.9 (61.9)	−206 (−161)	−0.03
TOPO	17.2	24.8 (22.4)	447 (478)	0.47	19.5***	64.2 (55.9)	−208 (−16)	−0.03
VEG	13.4	20 (20)	512 (510)	0.50	20.9***	63.9 (61.9)	−116 (−57)	−0.03
Structure								
STRUCT	14.5	29.5 (14.5)	381 (549)	0.30	5.7***	60 (40.1)	−103 (279)	−0.03
History								
KINGDOM	—	20.4 (13.1)	505 (605)	0.35	—	57.8 (39.7)	−33 (296)	−0.02
Others								
BIOME	—	46.4 (30.2)	120 (391)	0.18	—	62.4 (42.3)	−155 (248)	−0.02
LAT	−12.4	22.4 (13)	480 (595)	0.36	−5.6***	58.2 (40)	−57 (280)	−0.02

Species richness and all continuous predictor variables (except for VEG, TOPO, and TOPOVEG) were log<sub>10</sub>-transformed. Null model: deviance = 123.01; AIC = 737.61; *n* = 1,032. Values in parentheses refer to models without control for area. Because GLMs do not remove spatial autocorrelation from the residuals, significance levels are not reported. High percentage of explained deviance in single predictor SLM is mostly due to the strong influence of spatial trend term. SLMs leave no significant spatial autocorrelation in the residuals (all global Moran's *I* have *P* > 0.05). PRE, mean annual precipitation [millimeters per year (mm/a)]; WAT-ENER, water–energy model according to ref. 36; KINGDOM, floristic kingdom membership; BIOME, biome membership; LAT, absolute latitude. \*, *P* < 0.05; \*\*\*, *P* < 0.001.

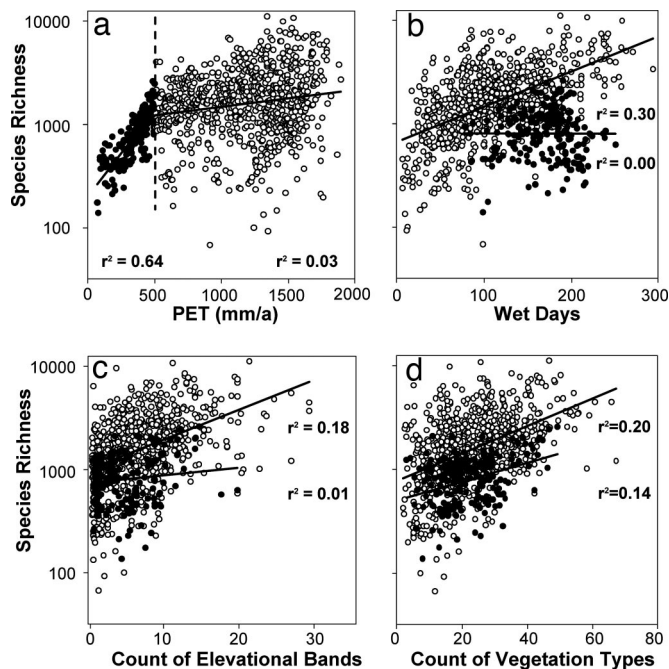
## Results and Discussion

The well known species–area relationship explains plant richness at local to regional scales (36). Interestingly, even across a variation of four orders of magnitude in our data set, area *per se* is a relatively weak predictor of species richness and explains only 6.6% of the global variation of plant species richness (Table 1). However, the explanatory power of area dramatically increases when spatial autocorrelation is explicitly modeled (57.4% deviance). This finding indicates strong neighborhood effects, which are also observed for subsequent environmental predictors [Table 1 and supporting information (SI) Table 3]. Furthermore, we find that regional spatial heterogeneity (measured as the number of vegetation types, elevational belts, or especially as a variable combining both) is a strong predictor of plant richness and is able to account for the effect of area (Table 1).

Among individual climatic variables, average annual temperature is thought to be of particular importance for ectothermic clades, given its exponential effect on rates of energy flux and thus, potentially, rates of biological interaction and diversification (3, 37). Although there is a significant positive effect of average annual temperature (TEMP) on vascular plant species richness (8.5% deviance; Table 1), there are indications for a quadratic rather than linear trend [ $\Delta$ AIC = 25; where  $\Delta$ AIC indicates the difference between the Akaike information criterion (AIC) of the model of interest and the AIC of the best fitting

model]. Under the generally untested assumption of uniform total abundance of individuals across space, the species–richness extension of the metabolic theory of ecology predicts a slope of 9.0 between the inverse of temperature (1,000/K) and the natural logarithm of species richness (37), which is very different from the one observed here [ $\ln(\text{richness}) = 13.88 - 1.89 \text{ temperature}^{-1}(1,000/\text{K})$ ; test for difference in slope: *t* = 54.33; *P* < 10<sup>−15</sup>].

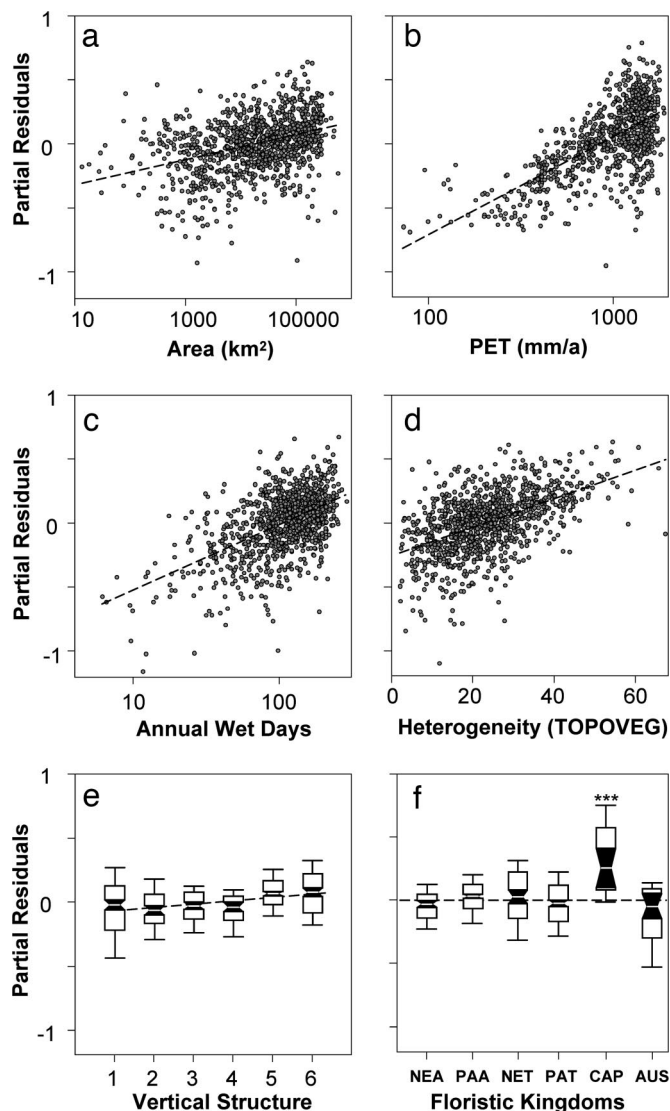
Actual evapotranspiration emerges as the strongest single climatic predictor [28.6% deviance generalized linear model (GLM); Table 1]. Mean annual net primary productivity yields somewhat poorer fits than actual evapotranspiration (26.7%;  $\Delta$ AIC = 28). Water–energy models that include interaction terms tend to have stronger explanatory power than those with only main effects confirming the important interdependence of these determinants. Of all potential combinations of energy-related variables and water-related variables, the full interaction model including potential evapotranspiration (PET) and the annual number of days with rainfall (WETDAYS, a variable that encapsulates both amount and temporal occurrence of precipitation) is the strongest (36.8% deviance). Other variable combinations to quantify the water–energy interaction (27, 38) yield significantly poorer fits ( $\Delta$ AIC > 44). Visual inspection and split-line analyses of the relationship between PET and species richness indicate a threshold at 505-mm PET, above which the



**Fig. 1.** Relationship between environmental predictors and species richness of vascular plants in low- and high-energy regions. Species richness is standardized to 10,000 km<sup>2</sup>. (a) The effect of PET [millimeters per year (mm/a)] on species richness. A close association is observed in regions with <505 mm/a PET (filled circles), whereas in regions with higher energy input (open circles) the relationship is not significant (breakpoint confirmed by split-line regression). (b–d) Also shown are relationships for wet days (b), topographical complexity measured as the number of elevational bands (c), and heterogeneity measured in number of vegetation types per region (d).

relationship between species richness becomes largely independent from further increasing annual energy input (Fig. 1a). Above the same breakpoint, the number of wet days, a nonsignificant predictor in low-energy regions, assumes strong predictive power (Fig. 1b), highlighting that water constrains richness only in high-energy regions (4). A similar interaction occurs with topographical complexity (Fig. 1c), but not habitat heterogeneity as such (Fig. 1d). The interaction with topography is qualitatively similar to previous findings for North American mammals (39) but with a breakpoint at much lower energy levels (505- vs. 1,000-mm PET). This finding strongly points to the different energetic and physiological constraints between these two groups (ectothermic vs. endothermic). These differential constraints may have affected rates of diversification and range limits of species. Determinants of species richness might change with spatial scale (31), and we therefore test for interactions between area and all predictor variables. Despite the large variation in areas, no significant effects emerge ( $\Delta$ AIC between 0.3 and 2.7), corroborating the validity of our model results across a wide range of scales.

Historical effects, i.e., regional differences in rates of past speciation, extinction, and dispersal, are notoriously difficult to quantify and often covary with contemporary environment and physiography. Topographic heterogeneity often is associated with a high potential for speciation during past periods of climate change (25, 34, 40) or during a recent uplift of mountain ranges, such as the Andes or Himalaya (25, 41, 42). We already noted a strong effect of topographic complexity. An alternative way to capture regional histories is to compare historically distinct biotas composed of almost completely nonoverlapping taxa—i.e., realms or, in our case, six floristic kingdoms (KINGDOM; largely following ref. 43), which may be considered as statistical



**Fig. 2.** Partial residuals plots for all variables included in the combined model of global plant richness (compare Table 2). These plots show the effects of a given variable when all others in the model are statistically controlled for. (a–e) Hatched lines partial fits. (e and f) Boxes indicate second and third quartiles, black notches denote 95% confidence intervals, and whiskers indicate 10th and 90th percentiles. NEA, Nearctic; PAA, Palaearctic; NET, Neotropic; PAT, Paleotropic; CAP, Cape; AUS, Australis. Note high partial residuals of the Cape floristic kingdom after controlled-for environmental differences (\*\*\*, significant at  $P < 0.001$ ; Tukey post hoc test). Specifically, a partial residual plot is a plot of  $r_i + b_k \times i_k$  vs.  $x_{ik}$ , where  $r_i$  is the ordinary residual for the  $i$ th observation,  $x_{ik}$  is the  $i$ th observation of the  $k$ th predictor, and  $b_k$  is the regression coefficient estimate for the  $k$ th predictor.

replicates. Kingdom membership alone explains a substantial amount of deviance (13.1%), a value that decreases to 2.9% when kingdoms are combined into three broad longitudinal bands (Americas, Europe–Africa, Asia–Australia) to minimize environmental collinearity.

We proceed to select the best single predictor variable or interaction term from each category to construct a combined multipredictor model. Consisting of six explanatory variables (AREA, PET, WETDAYS, TOPOVEG, STRUCT, and KINGDOM), the combined model explains 65.9% of the observed deviance in a GLM framework ( $AIC = -353.5$ ) and 70.2% in spatial linear model (SLM) ( $AIC = -456.9$ ) (Fig. 2 and SI Fig. 4). The SLM approach successfully removes spatial autocorre-

**Table 2. Global model of plant diversity**

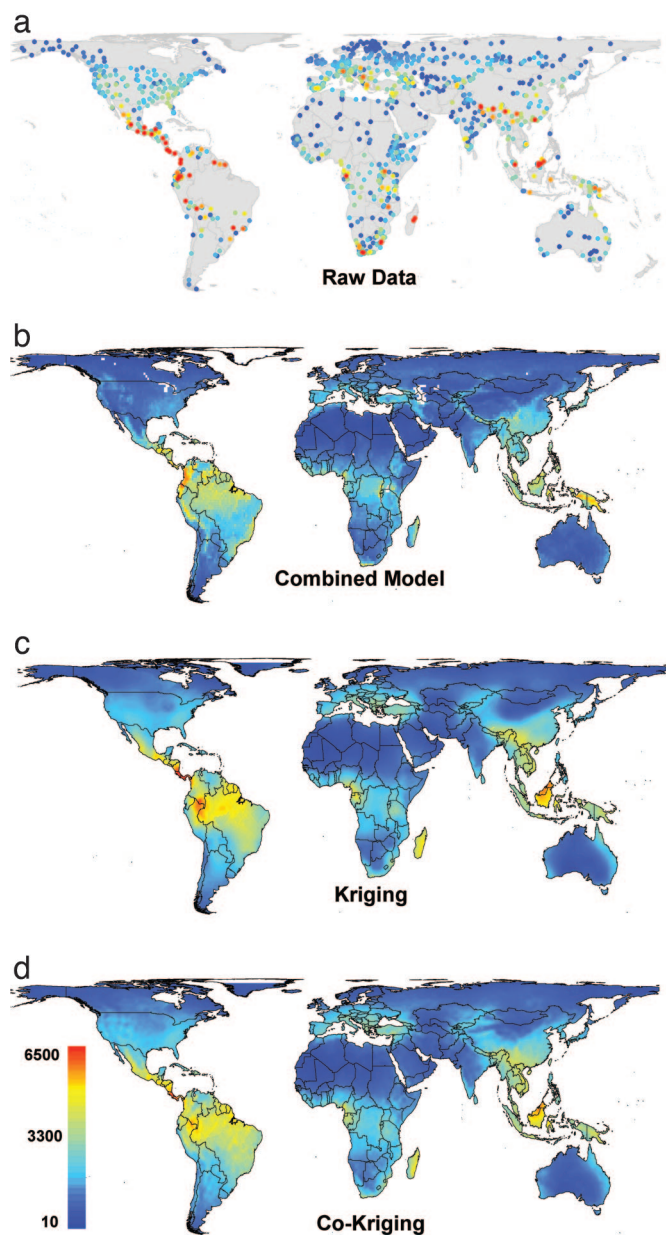
Combined model	GLM		SLM	
	Coefficient	t	Coefficient	z
AREA	0.096	9.4	0.118	11.5***
PET	0.759	18.2	0.747	12.4***
WETDAYS	0.507	14.9	0.542	12.3***
TOPOVEG	0.011	14.9	0.010	11.3***
STRUCT	0.030	5.9	0.022	4.5***
KINGDOM				
NEA	-0.154	-2.2	-0.081	-1.7
AUS	-0.061	-3.9	-0.162	-2.2*
CAP	0.285	6.1	0.281	4.1***
PAT	-0.051	-2.3	-0.062	-1.5
PAA	-0.006	-0.2	-0.023	-0.5
Deviance, %	65.9		70.2	
AIC	-353.5		-456.9	
Moran's I	0.17***		-0.01NS	

Results of GLM and SLM of a combined six-predictor model. KINGDOM: NEA, Nearctic; AUS, Australis; CAP, Capensis; PAT, Palearctic; PAA, Palaeartic. Estimates for KINGDOM refer to deviations from Neotropic (NET). NS, not significant; \*\*\*,  $P < 0.001$ .

lation from the model residuals (SI Fig. 5; for a map of GLM and SLM residuals, see SI Fig. 6). No interaction terms (including the previously asserted energy–water interactions) significantly improve the model fit. The most important model predictor is PET (11.1% partial deviance; Fig. 2) followed by number of wet days and environmental heterogeneity (both 7.2%).

When controlling for environmental dissimilarities in the combined model, floristic kingdom only has a small, yet significant, effect on richness [combined model (Table 2) compared with a model without term for kingdom:  $\Delta AIC = 68$  and 2.7% deviance, GLM;  $\Delta AIC = 29$  and 0.9% deviance, SLM]. Despite above and beyond differences due to environment, the world's floristic kingdoms appear to be remarkably similar in richness. There is one glaring exception: the Southern African Cape region, highlighted before for its unique biota and apparent high richness (16, 44) but never evaluated in the global context. We find that translated into species numbers, the Cape flora has more than twice as many species (on average 655 species per  $\approx 12,100 \text{ km}^2$  grid cell more; maximum: 1,637) per unit area than expected given its contemporary environment and topography, confirming, from a global perspective, its outstanding richness (44). The potential causes of the unique plant diversity of the Cape region are still debated and include climatic shifts from summer to winter rains starting in the Oligocene, pollinator specialization, mesoscale habitat specialization, and fire regimes, giving rise to an enormous diversification in some clades (45, 45). Crucially, we find that many regional differences in species richness that have classically been attributed to historical factors can also be predicted by contemporary differences in the environment. For instance, the long recognized greater diversity of Neotropical rainforests in relation to their African counterparts (mean  $\pm 95\%$  confidence interval; number of species per  $12,100 \text{ km}^2$ :  $2,479 \pm 39$  vs.  $1,886 \pm 43$ ) can at least statistically be predicted by environmental differences alone (e.g., mean annual precipitation:  $2,186 \pm 49$  vs.  $1,661 \pm 57$  mm; mean wet days:  $199 \pm 3$  vs.  $133 \pm 3$ ).

In summary, our combined model successfully explains the latitudinal gradient of plant species richness (SI Fig. 4), and the predicted global map (Fig. 3b) confirms many regional trends and hotspots anticipated before (11, 13, 14, 16, 17, 46). Geostatistical models (Fig. 3c and d) additionally incorporate information of environmental covariation and neighborhood effects. Given the importance of these effects asserted in the SLM, they



**Fig. 3.** Global patterns of vascular plant species richness. (a) The geographic distribution of the richness data of vascular plants for the 1,032 geographic regions analyzed in this study (each dot presents the mass centroid of a geographic entity; note that regions differ in size and that species counts have not been standardized). (b–d) The species-richness maps show area-standardized predictions of three different global models across an equal area grid ( $\approx 12,100 \text{ km}^2$ ,  $\approx 1^\circ$  latitude  $\times 1^\circ$  longitude near the equator) based on the combined multipredictor model (b), ordinary kriging of species richness (where species richness is interpolated purely as a function of spatial autocorrelation in the response variable) (c), and ordinary cokriging (which incorporates both the spatial autocorrelation in species richness and the combined model as an underlying trend) (d).

improve the quality of predictions, especially in relatively well sampled regions (e.g., North America, Europe, and Mesoamerica).

We have shown that relatively few variables, namely a combination of high annual energy input with constant water supply and extraordinarily high spatiotopographic complexity, are able to accurately predict the location of global centers of plant richness (Costa Rica–Chocó, Tropical Eastern Andes, Atlantic

Brazil, Northern Borneo, and New Guinea) (14, 17). Astonishingly, the most species-rich grid cell according to our combined model (situated between Colombia and Ecuador; Fig. 3*b*) has already been identified as such by Alexander von Humboldt (47) 200 years ago in the parlance of his time:

This portion of the surface of the globe affords in the smallest space the greatest possible variety of impressions from the contemplation of nature. . . . There, the different climates are ranged the one above the other, stage by stage, like the vegetable zones, whose succession they limit; and there the observer may readily trace the laws that regulate the diminution of heat, as they stand indelibly inscribed on the rocky walls and abrupt declivities of the Cordilleras.

Over the last decades, much controversy has arisen from the ambition to find one single factor that explains the enigmatic latitudinal gradient in species richness. Our findings demonstrate that different hypotheses are not mutually exclusive and core drivers likely act synergistically. Furthermore, our findings illustrate the significant advance that spatial analysis techniques applied to even megadiverse taxa at the global scale allow to both improve conceptual understanding as well as the quantitative knowledge base for conservation of yet understudied taxa. The challenge now is to close the many alarming taxonomic and geographic gaps in data availability that, surely to Humboldt's dismay, still exist.

## Materials and Methods

**Richness Data.** We analyzed the species richness of vascular plants (i.e., ferns, gymnosperms, and angiosperms) across 1,032 geographic units worldwide (Fig. 3*a*). Geographic units represent natural (e.g., mountain ranges, desert, and biogeographic provinces) or political units (e.g., countries, provinces, and national parks) and were derived from floras, checklists, and other literature sources (refer to ref. 46 for a full list of references). The data set was originally assembled to produce expert opinion-based global maps of plant species richness (11, 14, 17, 19, 46) and continental geostatistical (16, 48, 49) analyses. The original data set consists of >3,300 species-richness accounts referring to >1,800 geographic units. We excluded oceanic islands because isolation or geological age play a major role in the assembly of island floras (50). Furthermore, we excluded geographic units with an area of <10 km<sup>2</sup> and >300,000 km<sup>2</sup> to avoid spatially overlapping units. The data set covers almost the full spectrum of the global variation in abiotic conditions and includes all major biomes and floristic kingdoms.

**Putative Determinants.** We tested 40 variables as potential determinants of species richness (see SI Table 3 for full descriptions and references of all examined variables). Climatic variables and net primary productivity data were derived from interpolated, digitally available global data sets. The net primary productivity data set represents an average of 17 different global models (see ref. 51 for details). Mean values were extracted across all 1,032 investigated geographical units in ArcINFO (ESRI, Redlands, CA). Water–energy dynamics received particular attention in our analyses. We analyzed the predictive power of variables or variable combinations that have been previously reported to be strong predictors of plant richness (4, 27, 38) as well as all other possible combinations of energy-related and water-related variables. To analyze the potential effects of historical contingencies, we included floristic kingdom membership as a further variable (43). When simultaneously controlling for environmental dissimilarity of core predictors, deviations in species richness from the global environmental trend may point to an additional influence of idiosyncratic historical events on species richness

(40, 52). We also tested different variables describing habitat heterogeneity. The number of 300-m elevational belts per geographic unit (range of elevation divided by 300; TOPO) was calculated as a proxy of topographical complexity (31) by using the GTOPO-30 digital elevation model. Furthermore, the number of different vegetation types (VEG) and soil types (SOIL) occurring in a geographic unit were counted. We additionally created the combined variable TOPOVEG (TOPO + VEG) because global land cover data tend to underestimate changes along elevational gradients. As a measure of space use of the vegetation, all biomes were ranked according to their three-dimensional structural complexity (STRUCT). This measure varies from one (desert and tundra) to six (tropical broadleaf forest) along an integer scale. Variables were assigned to different categories corresponding to different hypotheses of species-richness gradients.

**GLM Analyses.** First, we performed GLMs to analyze potential single predictors of species richness. In a second step, more complex models were created (compare Tables 1 and 2 and SI Table 4). The fit of individual models is reported by using the proportion of deviance explained [deviance =  $-2 \times$  maximized log-likelihood; percentage deviance explained =  $(100 - \text{null deviance} / \text{residual deviance}) \times 100$ ], because that makes GLM and SLM results directly comparable. The best single predictor or combination of predictors from each category was included into a combined multipredictor model. The goodness-of-fit in relation to the model complexity was evaluated by using the AIC, which incorporates the maximized log-likelihood of the model and a term that penalizes models with greater complexity (53). Model selection was then based on  $\Delta\text{AIC}$ , which is the difference between the AIC of the model of interest and the AIC of the best fitting model (53).

**Spatial Analyses.** Spatial autocorrelation of species richness and predictor variables is a general feature of macroecological data sets (54). It inflates type I errors of traditional statistical tests and might affect parameter estimates (55). Because spatial autocorrelation also is present in the data set analyzed here, we performed simultaneous autoregressive models. Three different simultaneous autoregressive model types (lagged-response, lagged-mixed, and spatial error) were evaluated with different neighborhood structures and spatial weights (lag distances between 200 and 2,000 km, weighted and binary neighborhood coding). Final model selection was based on the reduction of spatial autocorrelation in the residuals and the minimization of AIC values. Simultaneous autoregressive models of the spatial error model type with a lag distance of 800 km and weighted neighborhood structure accounted best for the spatial structure in the analyzed data set. Spatial statistics were performed with the “spdep” library in the R software package (56). We assessed spatial autocorrelation in model residuals by using Moran's *I*, which can be considered a spatial equivalent to Pearson's correlation coefficient and normally varies between 1 (positive autocorrelation) and  $-1$  (negative autocorrelation). The expected Moran's *I* value for lacking spatial autocorrelation is close to 0 (57). Spatial Moran's *I* correlograms for the response variable as well as for the GLM and SLM residuals are provided in SI Fig. 5.

**Global Predictions.** We derive global predictions of plant richness across an equal area grid of  $\approx 110 \times 110$  km (12,100 km<sup>2</sup>; approximating an area of 1° latitude  $\times$  1° longitude near the equator) (Fig. 3*b*). We first make predictions based on the parameter estimates of the combined GLM and by using the same predictor variables for the global grid. This approach does not account for the spatial structure in the data except for what is dictated by the predictor variables. In the absence of richness

data for unsampled neighboring locations, the incorporation of the spatial autocorrelation signal when predicting into unsampled areas is not trivial. We used a two-level approach. First, we applied the geostatistical interpolation technique of ordinary kriging, which is commonly applied in other disciplines, such as mining, meteorology, and soil research (57). This approach interpolates between sampled quadrats exclusively according to the spatial dependence of the response variable and ignores underlying environmental gradients but has the advantage of an exact interpolation method at sampled locations (Fig. 3c). Second, we link this limited approach with the GLM-based environmental model by using ordinary cokriging, a commonly applied technique to enhance interpolation estimates (57). Whereas the former considers only the spatial dependence of the

response variable, the latter also accounts for the environmental covariation. The resulting global species-richness map (Fig. 3d) accounts for both environmental gradients and underlying spatial trends in the richness of plants. Geostatistical analyses were performed with the Geostatistical Analyst extension in ArcGIS (ESRI).

We thank Jens Mutke, Gerold Kier, and Wilhelm Barthlott (all of University of Bonn) for access to their global plant-richness data set; two anonymous reviewers for valuable comments; and Robert Ricklefs, David Currie, and Lauren Buckley for feedback on preliminary results. This study is part of the doctoral thesis of H.K., supervised by Wilhelm Barthlott at the University of Bonn. H.K. received funding and travel support from the German National Academic Foundation (Bonn, Germany).

- Schall JJ, Pianka ER (1978) *Science* 201:679–686.
- Wright DH (1983) *Oikos* 41:496–506.
- Rohde K (1992) *Oikos* 65:514–527.
- Hawkins BA, Field R, Cornell HV, Currie DJ, Guégan J-F, Kaufman DM, Kerr JT, Mittelbach GG, Oberdorff T, O'Brien EM, et al. (2003) *Ecology* 84:3105–3117.
- Currie DJ, Mittelbach G, Cornell HV, Field R, Guégan J, Hawkins BA, Kaufman DM, Kerr JT, Oberdorff T, O'Brien EM, Turner JRG (2004) *Ecol Lett* 7:1121–1134.
- Ricklefs RE (2004) *Ecol Lett* 7:1–15.
- Jetz W, Rahbek C (2002) *Science* 297:1548–1551.
- Lomolino MV (2004) in *Frontiers of Biogeography: New Directions in the Geography of Nature*, eds Lomolino MV, Heaney LR (Sinauer, Sunderland, MA), pp 293–296.
- Hutchinson GE (1959) *Am Nat* 93:145–159.
- Gaston KJ (1992) *Funct Ecol* 6:243–247.
- Mutke J, Barthlott W (2005) *Biologische Skrifter* 55:521–538.
- Currie DJ (1991) *Am Nat* 137:27–49.
- Myers N, Mittermeier RA, Mittermeier CG, da Fonseca GAB, Kent J (2000) *Nature* 403:853–858.
- Barthlott W, Lauer W, Placke A (1996) *Erdkunde* 50:317–328.
- Linder HP (1998) in *Chorology, Taxonomy and Ecology of the Floras of Africa and Madagascar*, eds Huxley CR, Lock JM, Cutler DF (R Botanic Gard, Kew, UK), pp 67–86.
- Mutke J, Kier G, Braun G, Schultz C, Barthlott W (2002) *Syst Geogr Plants* 71:1125–1136.
- Barthlott W, Mutke J, Rafiqpoor MD, Kier G, Kreft H (2005) *Nova Acta Leopoldina* 92:61–83.
- Davis SD, Heywood VH, Hamilton AC, eds (1994) *Centres of Plant Diversity: A Guide and Strategy for Their Conservation* (Int Union Conserv Nat Natural Resour, Cambridge, UK), Vol 1.
- Davis SD, Heywood VH, Hamilton AC, eds (1995) *Centres of Plant Diversity: A Guide and Strategy for Their Conservation* (Int Union Conserv Nat Natural Resour, Cambridge, UK), Vol 2.
- Davis SD, Heywood VH, Herrera-MacBryde O, Villa-Lobos J, Hamilton AC, eds (1997) *Centres of Plant Diversity: A Guide and Strategy for Their Conservation* (Int Union Conserv Nat Natural Resour, Cambridge, UK), Vol 3.
- Barthlott W, Biedinger N, Braun G, Feig F, Kier G, Mutke J (1999) *Acta Botanica Fennica* 162:103–110.
- Kier G, Küper W, Mutke J, Rafiqpoor MD, Barthlott W (2006) in *Taxonomy and Ecology of African Plants, Their Conservation and Sustainable Use*, eds Ghazanfar SA, Beentje HJ (Royal Botanic Gard, Kew, UK), pp 409–425.
- Lovett JC, Rudd S, Taplin J, Fridolf-Møller C (2000) *Biodivers Conserv* 9:37–46.
- Crisp MD, Laffan S, Linder HP, Monro A (2001) *J Biogeogr* 28:183–198.
- Kreft H, Sommer JH, Barthlott W (2006) *Ecography* 29:21–30.
- Wright DH, Currie DJ, Maurer BA (1993) in *Species Diversity in Ecological Communities: Historical and Geographical Perspectives*, eds Ricklefs RE, Schluter D (Univ of Chicago Press, Chicago), pp 66–74.
- Francis AP, Currie DJ (2003) *Am Nat* 161:523–536.
- Field R, O'Brien EM, Whittaker RJ (2005) *Ecology* 86:2263–2277.
- Wiens JJ, Donoghue MJ (2004) *Trends Ecol Evol* 19:639–644.
- Shmida A, Wilson MV (1985) *J Biogeogr* 12:1–20.
- Rahbek C, Graves GR (2001) *Proc Natl Acad Sci USA* 98:4534–4539.
- Ricklefs RE (1987) *Science* 235:167–171.
- Dynesius M, Jansson R (2000) *Proc Natl Acad Sci USA* 97:9115–9120.
- Jetz W, Rahbek C, Colwell RK (2004) *Ecol Lett* 7:1180–1191.
- Qian H, Ricklefs RE (2004) *Am Nat* 163:773–779.
- Rosenzweig ML (1995) *Species Diversity in Space and Time* (Cambridge Univ Press, Cambridge, UK).
- Allen AP, Brown JH, Gillooly J (2002) *Science* 297:1545–1548.
- Venevsky S, Veneskaia I (2003) *Ecol Lett* 6:1004–1016.
- Kerr JT, Packer L (1997) *Nature* 385:252–254.
- Qian H, Ricklefs RE (2000) *Nature* 407:180–182.
- Gentry AH (1982) *Ann Mo Bot Gard* 69:557–593.
- Hughes C, Eastwood R (2006) *Proc Natl Acad Sci USA* 103:10334–10339.
- Good R (1974) *The Geography of the Flowering Plants* (Longman, London).
- Linder HP (2003) *Biol Rev* 78:597–638.
- Cowling RM, Pressey RL (2001) *Proc Natl Acad Sci USA* 98:5452–5457.
- Kier G, Mutke J, Dinerstein E, Ricketts TH, Küper W, Kreft H, Barthlott W (2005) *J Biogeogr* 32:1107–1116.
- von Humboldt A (1845–1858) *Kosmos: Entwurf einer physischen Weltbeschreibung* (Cotta, Stuttgart/Tübingen, Germany).
- Mutke J (2002) *Räumliche Muster Biologischer Vielfalt: Die Gefäßpflanzenflora Amerikas im Globalen Kontext* (Univ of Bonn, Bonn).
- Mutke J, Barthlott W (2000) in *Results of Worldwide Ecological Studies. Proceedings of the First Symposium by the A.F.W. Schimper-Foundation*, eds Breckle S-W, Schweizer B, Arndt U (Heimbach, Stuttgart, Germany), pp 435–447.
- Whittaker RJ (1998) *Island Biogeography: Ecology, Evolution, and Conservation* (Oxford Univ Press, Oxford).
- Cramer W, Kicklighter DW, Bondeau A, Moore B, III, Churkina G, Nemry B, Ruimy A, Schloss AL, Participants of "Potsdam '95" (1999) *Global Change Biol* 5:1–15.
- Whittaker RJ, Willis KJ, Field R (2001) *J Biogeogr* 28:453–470.
- Johnson JB, Omland KS (2004) *Trends Ecol Evol* 19:101–108.
- Legendre P (1993) *Ecology* 74:1659–1673.
- Lennon JJ (2000) *Ecography* 23:101–113.
- R Development Core Team (2005) *R: A Language and Environment for Statistical Computing* (R Found Stat Comput, Vienna).
- Fortin M-J, Dale MRT (2005) *Spatial Analysis: A Guide for Ecologists* (Cambridge Univ Press, Cambridge, UK).