# Patterns, determinants and models of woody plant diversity in China

**Zhiheng Wang**[*,†], **Jingyun Fang, Zhiyao Tang and Xin Lin**

*Department of Ecology, College of Urban and Environmental Sciences, and Key Laboratory for Earth Surface Processes of the Ministry of Education, Peking University, Beijing 100871, China*

What determines large-scale patterns of species richness remains one of the most controversial issues in ecology. Using the distribution maps of 11 405 woody species in China, we compared the effects of habitat heterogeneity, human activities and different aspects of climate, particularly environmental energy, water–energy dynamics and winter frost, and explored how biogeographic affinities (tropical versus temperate) influence richness–climate relationships. We found that the species richness of trees, shrubs, lianas and all woody plants strongly correlated with each other, and more strongly correlated with the species richness of tropical affinity than with that of temperate affinity. The mean temperature of the coldest quarter was the strongest predictor of species richness, and its explanatory power for species richness was significantly higher for tropical affinity than for temperate affinity. These results suggest that the patterns of woody species richness mainly result from the increasing intensity of frost filtering for tropical species from the equator/lowlands towards the poles/highlands, and hence support the freezing-tolerance hypothesis. A model based on these results was developed, which explained 76–85% of species richness variation in China, and reasonably predicted the species richness of woody plants in North America and the Northern Hemisphere.

## 1. INTRODUCTION

The mechanism underlying the large-scale patterns of species richness is one of the most controversial issues in ecology [1]. In the past two decades, with the increasing availability of large-scale range maps of animals, considerable progress has been made for the continental and global patterns in species richness of mammals [2,3], birds [4–6] and amphibians [7]. Although vascular plants are one of the most important components of terrestrial ecosystems, continental and global patterns of plant richness have been poorly investigated largely owing to the lack of precise large-scale range maps [8].

To understand the mechanisms underlying the large-scale patterns of species richness, many hypotheses focusing on different aspects of contemporary climate have been proposed [9]. For example, two energy hypotheses state that species richness is primarily determined by energy availability, but focus on different energy variables (i.e. thermal versus chemical energy) [10]: (i) the species richness-productivity hypothesis, where energy is usually measured by net primary productivity (NPP) or annual evapotranspiration (AET) [11–13], and (ii) the ambient energy hypothesis, where energy is measured by mean annual temperature or potential evapotranspiration (PET) [2,14]. By contrast, the water–energy dynamics hypothesis [15] proposed that species richness is determined by the combined effects of available water (measured by rainfall or water deficit (WD) in linear form) and environmental energy (measured by minimum monthly or annual PET in parabolic form) [16,17]. Two global models based on this hypothesis have been proposed: the Interim General Model (including IGM-1 and IGM-2), [15,17] and Francis & Currie's model [16] (F&C's model for short). Besides, some other studies indicated that climatic seasonality influences species richness patterns by altering the allocation of energy use of individuals [18] or the length of growing season for plants [19].

Alternatively, according to freezing-tolerance hypothesis (or tropical conservatism hypothesis), species richness is primarily determined by winter coldness because most clades evolved in tropical-like climate and hence could hardly disperse into cold, temperate regions owing to their niche conservatism [20–22]. This hypothesis integrates the effects of contemporary climate with evolutionary history [22], which is one of the biggest challenges facing ecologists. Given this hypothesis, we can predict that winter coldness should affect the richness of the species with tropical affinities more strongly than those with temperate affinities. However, this prediction remains poorly tested.

Another set of hypotheses predict that habitat heterogeneity determines the patterns of species richness by its influence on species turnover [23] and (or) species diversification rates [24]. Habitat heterogeneity in a region is generally represented by topographic relief (e.g. altitude range) [23] and local climatic heterogeneity (i.e. the ranges of mean annual temperature (MAT) and mean annual precipitation (MAP)) [25]. Recently, rapidly growing intensity of human activities has become a potential driver of the large-scale patterns of species richness [26,27], and is increasingly attracting the attention of ecologists.

* Author for correspondence (zhiheng.wang@pku.edu.cn).
† Present address: Center for Macroecology, Evolution and Climate, Department of Biology, University of Copenhagen, Universitetsparken 15, DK-2100 Copenhagen Ø, Denmark.

In the North Hemisphere, China harbours a much richer flora and also a steeper species richness gradient than North America and Europe, which makes China immensely suitable for testing the hypotheses explaining the large-scale patterns of species richness. Here, using the distribution maps of woody plants in China [28,29], we: (i) explored the geographical patterns in species richness of all woody plants, trees, shrubs and woody lianas and their concordance, and identified the primary determinant for the richness patterns by comparing between the effects of factors representing different hypotheses; (ii) evaluated the effects of biogeographic affinities on species richness patterns and their relationships with environmental factors; and (iii) developed models and used them to predict the species richness patterns of woody plants in the Northern Hemisphere.

## 2. DATA AND METHODS

### (a) *Species richness of woody plants*

The species distribution maps were from the *Database of China's Woody Plants* (http://www.ecology.pku.edu.cn/plants/woody/index.asp) [28,29], which contains 11 405 native woody species, including 3165 trees, 7205 shrubs and 1035 lianas (see the electronic supplementary material, table S1). Exotic species were excluded from the database. The taxonomy of this database was updated following the recently published *Flora of China* (http://www.efloras.org/) and Species2000 (Checklist 2008, http://www.sp2000.org/), where the taxonomy is current and comparable with that used in other regions. The species distributions in the database were compiled from all national-level floras published before 2008, including *Flora Reipublicae Popularis Sinicae* (126 issues of 80 volumes), *Flora of China* and *Higher Plants of China* (10 volumes) [30], more than 120 volumes of provincial floras, and a great number of local floras and inventory reports across the country. To improve the quality of species range maps in the database, 21 experts from different regions in China were invited to check and supplement the species distributions in every region.

The database provides the species distribution maps at two spatial resolutions: counties with a median area of 2081 km$^2$ (skewness = 9.93) and grids of $50 \times 50$ km. To eliminate the influence of area on the estimation of species richness [1], the maps based on equal-area grids were used. As the grids located on the borders or along coasts are usually incomplete, we excluded those with land area smaller than 1250 km$^2$. A total of 3794 grids were finally used in our analyses. Species richness was estimated for four species groups: all woody plants, trees, shrubs and lianas. As spatial scale can potentially influence the relationships between species richness and environmental factors [4], we repeated all the analyses using grids of $100 \times 100$ km, and found that the results at the two spatial scales were consistent. Therefore, we reported only the results for grids of $50 \times 50$ km.

To evaluate the effects of long-term evolution, we categorized the species within each species group into three biogeographic affinities. Based on the evolution of China's flora and its relationship with the floras of other major biogeographic regions, the regions where the genera and their families are believed to have diversified, and also the global distributions of the genera, Wu *et al.* [31] divided the

genera of China's vascular plants into three major biogeographic affinities: tropical, temperate and cosmopolitan [32]. Following Harrison & Grace [33], we defined the biogeographic affinity of a species as that of its genus (Harrison & Grace used family), and finally recognized 5682 (49.8%) tropical species, 4895 (42.9%) temperate species and 666 (5.8%) cosmopolitan species (electronic supplementary material, table S1). Then the species richness patterns of the three biogeographic affinities were estimated for all woody plants and the three lifeforms, respectively, and were compared with the overall species richness of each group. The influence of biogeographic affinities on the relationships between species richness and climatic factors was investigated. In our analyses, we focused on the comparisons between the richness patterns of tropical and temperate species.

### (b) *Environmental factors*

Climatic data with the resolution of 30 arc-second (*ca* 1 km at the equator) were obtained from the WorldClim website [34]. The database includes mean monthly temperature (MMT) and MAT (in °C) and mean monthly precipitation (MMP) and MAP (in mm), the mean temperature of the coldest quarter (MTCQ, in °C) and the mean temperature of the warmest quarter (MTWQ, in °C), the precipitation of the driest quarter (PDQ, in mm), the annual range of temperature (ART, in °C) and the seasonality of temperature (TSN, defined as the standard deviation of MMT) and precipitation (PSN, defined as the coefficient of variation of MMP).

Using MMT and MMP, the following variables were calculated: monthly/annual PET (mm) and AET (mm; calculated using the method of [35]), moisture index ($I_m$), WD (mm), warmth index (WI, °C) and annual rainfall (RAIN, mm). PET is widely used as a measure of ambient (or thermal) energy [2]. We used both annual and minimum monthly PET (PET and PET$_{min}$, respectively) to evaluate the water–energy dynamics hypothesis [15–17]. AET reflects the amount of water that plants can actually use, and is usually used as a surrogate of NPP [13]. $I_m$ represents the environmental humidity [35], whereas WD represents the aridity and is defined as the difference between PET and AET [16]. WI has been widely used to determine the distributions of species and vegetation in eastern Asia [36], and is defined as:

$$\text{WI} = \text{sum of } (\text{MMT} - 5), \quad \text{MMT} > 5°\text{C},$$

where MMT is mean monthly temperature. RAIN is the sum of the MMP when MMT > 0°C [15,19].

All climatic variables were grouped into three categories: (i) environmental energy, including MAT, MTCQ, MTWQ, WI, PET and PET$_{min}$; (ii) water availability, including MAP, PDQ, RAIN, $I_m$, AET and WD; and (iii) climatic seasonality, including ART, TSN and PSN. The value of a grid for each variable was estimated by averaging all cells in that grid. For comparison with previous studies [12,13], we included both MAP and RAIN in our analysis.

Three variables were used to estimate habitat heterogeneity: the ranges of altitude (TOPO), MAT (RMAT) and MAP (RMAP) within grids. Altitudinal range was calculated as the difference between the maximum and minimum elevations of a grid using a GTOPO30 digital
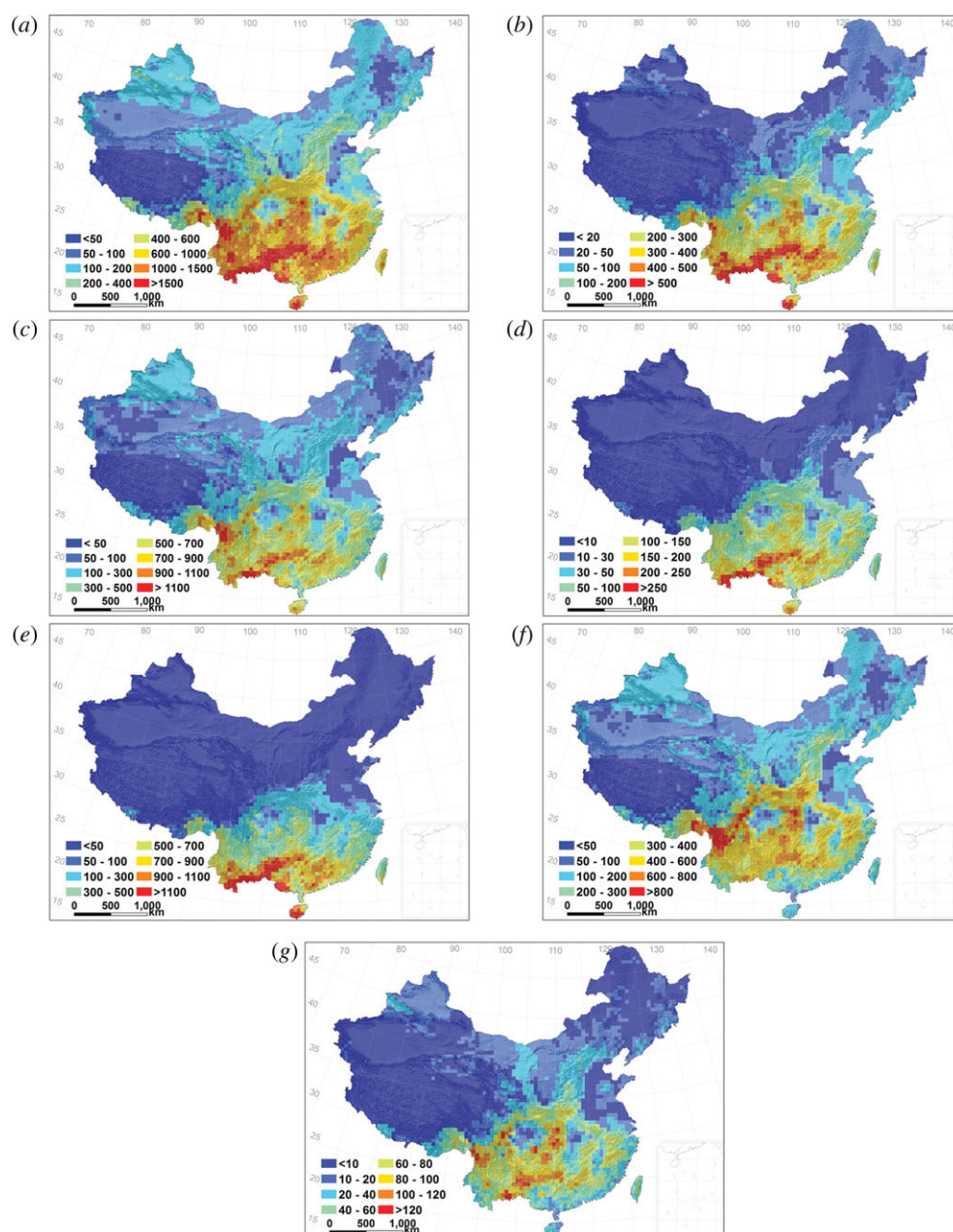
Figure 1. Geographical patterns in the species richness of woody plants in China estimated in equal-area grids of 50 × 50 km. (*a*) All woody plants; (*b*) trees; (*c*) shrubs; (*d*) woody lianas; (*e–g*) the species with tropical, temperate and cosmopolitan affinity, respectively.

elevational model, and was used to represent topographic relief. RMAT (or RMAP) was calculated as the difference between the maximum and minimum MAT (or MAP) in a grid, and was used to represent the heterogeneity of climatic conditions. For comparison with previous studies [15,17,19,37], these variables were log-transformed.

Finally, human population density (HPD), gross domestic product (GDP) and area of cropland per grid (CROP) were used to represent human activities. The average HPD and GDP of 2003–2005 in 2408 counties of China were from the *China Statistical Yearbook for Regional Economy* (2003–2005), and were interpolated into raster files of 5 × 5 km, which were used to calculate the mean HPD and GDP in each grid. CROP was extracted from the 1 : 1 000 000 vegetation atlas of China [38].

The statistics of the above variables for China and the Northern Hemisphere, and the correlation coefficients of

the variables against each other are presented in the electronic supplementary material, appendix S1. Most climatic variables in China cover *ca* three-quarters of their ranges in the entire Northern Hemisphere, suggesting that China is the proper representative for the Northern Hemisphere in terms of climate.

### (c) *Data analysis and model development*

Correlation analyses were conducted to evaluate the concordance between the species richness patterns of all woody plants and the three lifeforms, and also between the overall species richness and the richness of species with tropical, temperate and cosmopolitan affinities within each species group.

Bivariate regressions were used to evaluate the explanatory power of each predictor for the species richness of the

four species groups. Additionally, following the water–energy dynamics hypothesis [15–17,19], we also evaluated the explanatory power of the water–energy dynamics functions combining optimal energy with linear water variables (i.e. $PET_{min} - PET_{min}^2 + RAIN$ and $PET - PET^2 - WD$) using multiple regressions. Because species richness data usually satisfy Poisson distribution, generalized linear models (GLMs) with Poissonian residuals were used for all regression analyses [39]. The coefficients of determination ($r^2$) of the models were estimated as: $r^2 = 100 \times (1 - \text{residual variation/null variation})$. For multiple regressions, adjusted $r^2$ was used.

We conducted partial regressions [40] to compare the effects of climate with habitat heterogeneity and human activity. By partial regressions, the total variation of species richness was partitioned into: (i) independent components; (ii) covarying component; and (iii) unexplained variation.

Then, we developed combined models for the species richness of all woody plants and the three lifeforms using the following methods: (i) the best individual predictor for the species richness of a species group was kept in its model; (ii) to avoid the multi-collinearity between the predictors of the same environmental category, only one variable from each category was allowed to enter the model. Because the variables of human activities were not significant for most species groups (see §3), they were not included in the models; and (iii) all the possible combinations of predictors following the above criteria were examined, and the model with the lowest akaike information criterion (AIC) was selected for each species group. Adjusted $r^2$ was given for the selected models. Variance inflation factors (VIFs) were also calculated for all models to evaluate the significance of multi-collinearity [40]. Generally, if VIF is greater than five, the multi-collinearity is considered as significant.

We tested the combined models using the species richness of trees in North America. First, the tree species richness in grids of $50 \times 50$ km was estimated using an *Atlas of United States Trees* [28], and predicted by our model using the environmental data of North America. It is noteworthy that the predicting method of GLMs with Poisson residuals is different from those of ordinary least-square linear regressions [39]. Second, the predictions were plotted against the observations. The distances between the predictions and 1 : 1 line (a line with intercept = 0, slope = 1) were calculated, which represented the errors in the predictions and were used to evaluate the model performance: the distances of a better model should have a symmetric frequency distribution with the average closer to zero and smaller standard deviation.

Finally, using our model, we predicted the species richness patterns of all woody plants and the three lifeforms in the Northern Hemisphere.

For comparison, IGM2 [17] and F&C's model [16] were re-fitted using China's data, which were referred to as the water–energy dynamics model and refitted F&C's model, respectively, and were used to predict the species richness of North American trees.

Preliminary analyses indicated significant spatial autocorrelations in the raw data of species richness (electronic supplementary material, figure S1), which can inflate type I errors and hence the significance level of models and correlations because of the dependency in samples [41].

Table 1. Correlation coefficients between species richness patterns of different species groups and biogeographic affinities. (The *p*-values of all correlation coefficients tested using a bootstrap method (see §2) are smaller than 0.05.)

| | trees | shrubs | lianas |
|---|---|---|---|
| all woody plants | 0.99 | 0.99 | 0.97 |
| trees | — | 0.96 | 0.98 |
| shrubs | — | — | 0.93 |
| | tropical | temperate | cosmopolitan |
| all woody plants | | | |
|   overall | 0.93 | 0.87 | 0.93 |
|   tropical | — | 0.61 | 0.74 |
|   temperate | — | — | 0.96 |
| trees | | | |
|   overall | 0.94 | 0.88 | 0.65 |
|   tropical | — | 0.63 | 0.42 |
|   temperate | — | — | 0.80 |
| shrubs | | | |
|   overall | 0.91 | 0.85 | 0.94 |
|   tropical | — | 0.51 | 0.77 |
|   temperate | — | — | 0.87 |
| lianas | | | |
|   overall | 0.94 | 0.84 | 0.78 |
|   tropical | — | 0.50 | 0.46 |
|   temperate | — | — | 0.71 |

Therefore, in our analyses, we used a bootstrap method to test the significance of all correlation coefficients and models [41]. To do this, we first randomly re-sampled *ca* 8 per cent of all grids for each species richness variable to perform the correlation (or GLM) analysis, and this was repeated 1000 times. Only if more than 95 per cent of the repeats were significant, the correlation (or GLM) was considered as significant.

All statistical analyses were carried out using R (http://www.r-project.org/).

## 3. RESULTS

### (a) *Geographical patterns of species richness*
Species richness of all woody plants ranged from 2 to 2837, with an average of 358 species per grid (electronic supplementary material, table S1). The average species richness per grid for trees, shrubs and lianas was 104 (1–1028), 224 (2–1528) and 38 (1–327), respectively. The species richness of the four species groups was all highly right-skewed (skewness > 1.5, see the electronic supplementary material, table S1 and figure S2).

The patterns in the species richness of all woody plants, the three lifeforms and the three biogeographic affinities all matched the topographic structure in China: richness was high in mountains, but low in deserts, plains and basins (figure 1). The species richness patterns of all woody plants and the three lifeforms were highly concordant with each other ($r > 0.93$, $p < 0.05$; table 1), suggesting that they are plausibly determined by the same factors.

For the four species groups, the correlation coefficients between the overall and tropical species richness were 0.91–0.94, and were consistently higher than those between the overall and temperate species richness (0.84–0.88; table 1). Moreover, within each group, the richness patterns of temperate and cosmopolitan species

Table 2. Explanatory power ($r^2$) of the predictors for the species richness patterns of all woody plants, trees, shrubs and woody lianas evaluated by generalized linear models (GLM). (ALL, overall species richness; TROP, species with tropical affinities; TEMP, species with temperate affinities; COSM, species with cosmopolitan affinities. *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.)

| | all woody plants | | | | trees | | | | shrubs | | | | lianas | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | ALL | TROP | TEMP | COSM | ALL | TROP | TEMP | COSM | ALL | TROP | TEMP | COSM | ALL | TROP | TEMP | COSM |
| **environmental energy** | | | | | | | | | | | | | | | | |
| MAT | 53.5*** | 65.4*** | 28.7*** | 38.3*** | 58.9*** | 62.4*** | 37.4*** | 19.5*** | 46.2*** | 62.9*** | 17.5*** | 36.4*** | 63.8*** | 50.7*** | 43.1*** | 29.1*** |
| MTWQ | 20.9*** | 22.2*** | 10.9*** | 14.1*** | 24.5*** | 17.9*** | 16.3*** | 3.0 | 16.3*** | 19.9*** | 4.8 | 12.3*** | 29.4*** | 13.3*** | 21.8*** | 9.0*** |
| MTCQ | 66.0*** | 78.9*** | 37.6*** | 48.7*** | 69.4*** | 76.2*** | 44.7*** | 31.5*** | 60.1*** | 77.1*** | 26.2*** | 47.1*** | 73.0*** | 66.2*** | 48.6*** | 41.6*** |
| WI | 42.1*** | 55.1*** | 17.2*** | 25.7*** | 48.0*** | 53.0*** | 24.9*** | 9.2*** | 35.3*** | 52.0*** | 8.2*** | 24.2*** | 52.4*** | 45.6*** | 28.6*** | 17.8*** |
| PET | 37.0*** | 45.4*** | 15.7*** | 22.4*** | 41.1*** | 41.8*** | 22.3*** | 8.2*** | 30.8*** | 43.3*** | 7.7** | 20.3*** | 45.2*** | 34.8*** | 25.4*** | 16.0*** |
| PET$_{min}$ | 40.1*** | 52.3*** | 13.5*** | 21.6*** | 41.9*** | 53.1*** | 17.7*** | 11.7*** | 37.0*** | 50.5*** | 8.8*** | 21.6*** | 40.8*** | 51.4*** | 11.6*** | 17.8*** |
| **water availability** | | | | | | | | | | | | | | | | |
| MAP | 57.6*** | 57.3*** | 38.4*** | 45.3*** | 60.6*** | 54.8*** | 47.0*** | 30.6*** | 52.6*** | 54.7*** | 26.0*** | 44.2*** | 58.5*** | 31.1*** | 41.8*** | 39.0*** |
| RAIN | 58.6*** | 58.2*** | 38.7*** | 45.7*** | 61.1*** | 55.0*** | 47.4*** | 31.1*** | 53.5*** | 55.5*** | 26.2*** | 45.0*** | 59.4*** | 32.1*** | 42.3*** | 39.4*** |
| PDQ | 31.9*** | 29.7*** | 20.8*** | 26.7*** | 34.2*** | 26.5*** | 27.4*** | 17.2*** | 28.1*** | 27.9*** | 11.3*** | 27.0*** | 34.6*** | 8.2*** | 36.9*** | 20.0*** |
| AET | 57.3*** | 64.3*** | 33.9*** | 41.1*** | 64.0*** | 62.1*** | 47.1*** | 27.6*** | 49.6*** | 60.9*** | 20.0*** | 39.9*** | 64.3*** | 39.9*** | 38.5*** | 35.1*** |
| $I_m$ | 45.8*** | 37.3*** | 40.4*** | 41.0*** | 45.9*** | 35.8*** | 45.5*** | 35.3*** | 44.1*** | 35.2*** | 32.0*** | 39.9*** | 40.4*** | 9.1*** | 30.8*** | 42.4*** |
| WD | 28.4*** | 37.1*** | 21.7*** | 24.5*** | 35.1*** | 40.9*** | 34.0*** | 21.9*** | 24.4*** | 34.6*** | 14.2*** | 24.6*** | 35.9*** | 6.6* | 26.8*** | 28.0*** |
| **climatic seasonality** | | | | | | | | | | | | | | | | |
| ART | 61.9*** | 77.1*** | 35.8*** | 48.2*** | 63.7*** | 73.8*** | 41.2*** | 41.0*** | 58.3*** | 77.6*** | 28.1*** | 49.3*** | 66.2*** | 60.7*** | 43.4*** | 46.9*** |
| TSN | 48.6*** | 63.5*** | 27.2*** | 37.3*** | 49.0*** | 60.6*** | 29.7*** | 35.9*** | 47.1*** | 64.5*** | 23.7*** | 38.7*** | 49.8*** | 47.7*** | 33.2*** | 38.9*** |
| PSN | 24.8*** | 19.6*** | 20.4*** | 23.6*** | 23.9*** | 20.2*** | 20.6*** | 12.6*** | 23.0*** | 18.7*** | 15.1*** | 24.1*** | 34.9*** | 6.9*** | 51.0*** | 20.7*** |
| **habitat heterogeneity** | | | | | | | | | | | | | | | | |
| TOPO | 13.9*** | 8.0*** | 19.2*** | 16.5*** | 10.4*** | 7.4*** | 13.3*** | 12.4*** | 16.9*** | 9.5*** | 23.7*** | 18.1*** | 8.8*** | 5.2*** | 18.8*** | 16.5*** |
| RMAT | 12.8*** | 7.6*** | 17.6*** | 14.9*** | 9.5*** | 6.9*** | 12.3*** | 11.8*** | 15.6*** | 8.7*** | 22.0*** | 16.5*** | 7.8*** | 4.8*** | 16.7*** | 15.4*** |
| RMAP | 47.4*** | 37.5*** | 42.3*** | 41.3*** | 45.3*** | 34.9*** | 43.0*** | 26.6*** | 47.2*** | 35.4*** | 36.4*** | 40.5*** | 40.1*** | 11.6*** | 26.5*** | 38.2*** |
| **human activity** | | | | | | | | | | | | | | | | |
| HPD | 3.3 | 2.0 | 1.9 | 2.3 | 3.8* | 0.6 | 3.5* | 1.7 | 2.5 | 1.3 | 0.5 | 1.8 | 3.2 | 2.7 | 0.4 | 1.5 |
| GDP | 1.1 | 1.2 | 0.2 | 0.3 | 1.2 | 0.6 | 0.6 | 0.3 | 0.8 | 0.9 | 0.1 | 0.2 | 1.0 | 0.7 | 0.1 | 0.2 |
| CROP | 1.7 | 0.4 | 1.5 | 1.5 | 2.5 | 0.2 | 3.7* | 1.0 | 1.0 | 0.2 | 0.4 | 1.0 | 1.6 | 6.7*** | 0.8 | 1.0 |
| **water–energy dynamics** | | | | | | | | | | | | | | | | |
| PET$_{min}$ – PET$^2_{min}$ | 53.7*** | 67.2*** | 29.3*** | 39.6*** | 55.0*** | 66.1*** | 33.4*** | 25.5*** | 49.8*** | 65.8*** | 21.0*** | 40.3*** | 56.1*** | 61.7*** | 37.6*** | 30.0*** |
| PET$_{min}$ – PET$^2_{min}$ + RAIN | 62.4*** | 71.6*** | 41.8*** | 49.4*** | 64.9*** | 70.6*** | 49.8*** | 32.8*** | 57.3*** | 70.0*** | 29.1*** | 49.0*** | 64.7*** | 62.4*** | 51.8*** | 40.8*** |
| PET – PET$^2$ | 37.8*** | 47.6*** | 19.9*** | 25.6*** | 43.2*** | 43.9*** | 27.1*** | 9.1*** | 31.1*** | 45.0*** | 12.2*** | 23.4*** | 49.5*** | 36.1*** | 39.6*** | 18.0*** |
| PET – PET$^2$ – WD | 61.1*** | 68.5*** | 46.9*** | 51.0*** | 69.3*** | 65.9*** | 62.9*** | 33.2*** | 52.7*** | 65.0*** | 32.7*** | 48.7*** | 71.2*** | 41.5*** | 58.0*** | 45.1*** |

Table 3. Regression coefficients for the combined models developed in this study. ($p$-values for all models were $< 0.001$.)

| plant groups | coefficients | | | | | | adj. $r^2$ (%) | AIC |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| | intercept | MTCQ | WD | TSN | log (RMAT) | log (TOPO) | | |
| all woody plants | 4.2416 | 0.1033 | $-0.0016$ | 0.1255 | 0.5354 | — | 80.3 | 365 248 |
| trees | 2.5787 | 0.1242 | $-0.0027$ | 0.1752 | 0.5430 | — | 83.2 | 120 556 |
| shrubs | 1.5974 | 0.0848 | $-0.0011$ | 0.0847 | — | 0.4934 | 75.9 | 243 339 |
| | 4.0233 | 0.0877 | $-0.0011$ | 0.0952 | 0.5254 | — | 75.8 | 244 752 |
| liana | 1.2627 | 0.1466 | $-0.0037$ | 0.1874 | 0.5697 | — | 84.9 | 49 360 |
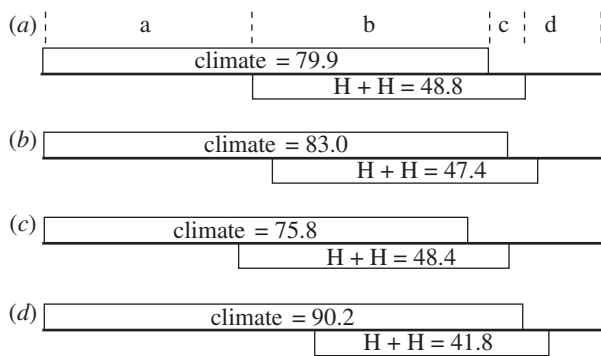


Figure 2. Comparisons between climatic effects and the effects of habitat heterogeneity and human activities on species richness by partial regressions. (*a*) all woody plants; (*b*) trees; (*c*) shrubs; (*d*) woody lianas. The variation of species richness is partitioned into (a) the independent components of climate and (c) habitat heterogeneity + human activities, (b) the covarying component and (d) residual variation. (*a*) a, 37.5; b, 42.4; c, 6.4; d, 13.7. (*b*) a, 40.8; b, 42.1; c, 5.3; d, 11.7. (*c*) a, 34.8; b, 41.0; c, 7.4; d, 16.9. (*d*) a, 48.4; b, 37.2; c, 4.6; d, 9.8.

were strongly correlated with each other, but both were moderately correlated with the richness of tropical species (table 1), which suggests that the determinants of tropical species richness may be different from those of temperate/cosmopolitan species richness.

### (b) Relationships between species richness and environmental variables

Among all variables, the MTCQ was consistently the strongest single predictor of species richness for all woody plants and the three lifeforms, and its $r^2$ was 60–73%, which was 10–15% higher than that of MAT, 28–35% higher than PET (table 2). Additionally, the $r^2$ of MTCQ was also 10–30% higher than the quadratic functions of PET and $PET_{min}$ and the combined water–energy dynamics functions despite their higher numbers of predictors (table 2). Annual rainfall was the best single water-related predictor for the species richness of all woody plants and shrubs, while AET was the best one for trees and lianas. The $r^2$ of annual rainfall was consistently higher than that of MAP, but the difference was small (table 2). The ART was the strongest predictor among the variables of climatic seasonality, while the range of annual precipitation was the strongest among the variables of habitat heterogeneity. By contrast, the variables of human activities were not significant for the species richness of most species groups (table 2).

Partial regressions indicated that climate independently accounted for 35–48% of richness variation after the effects of habitat heterogeneity and human activities were controlled. By contrast, habitat heterogeneity and human activities independently explained much less (5–7%) when climatic effects were controlled (figure 2).

The $r^2$ of the predictors was substantially different between biogeographic affinities (table 2). For all woody plants, trees and lianas, MTCQ was consistently the strongest predictor of the richness of tropical species, whereas RMAP, annual rainfall and precipitation seasonality were the strongest predictors for the richness of temperate species, respectively. For shrubs, the ART and MTCQ were the strongest predictors for tropical species, and the difference in the $r^2$ of the two variables was only 0.5 per cent. By contrast, RMAP was the strongest predictor for the richness of temperate shrub species.

### (c) Models for woody plant richness

The combined GLMs for the species richness of all woody plants and the three lifeforms selected consistent climatic variables (table 3): MTCQ, WD and temperature seasonality, representing the effects of environmental energy, water availability and climatic seasonality, respectively. As an indicator of habitat heterogeneity, RMAT was selected in the models for all woody plants, trees and lianas, while elevational range (TOPO) was selected in the model for shrubs. As RMAT and TOPO were strongly correlated ($r = 0.98$, $p < 0.001$; see the electronic supplementary material, appendix S1), replacing TOPO by RMAT in the model for shrubs reduced the $r^2$ only by 0.1 per cent (table 3). Therefore, we finally chose the model with RMAT for shrubs to get a consistent model for all woody plants and the three lifeforms:

$$\text{species richness} = a + b_1 \times \text{MTCQ} + b_2 \times \text{WD} + b_3 \times \text{TSN} + b_4 \times \text{RMAT},$$

where $a$, $b_1 - b_4$ were regression coefficients. The VIFs for the four variables in all models were smaller than five (electronic supplementary material, table S2), indicating insignificant multi-collinearity in the models.

The models reasonably predicted the species richness of China's woody plants, and the $r^2$ was 76–85% for all woody plants and the three lifeforms, which was 14–23% higher than those of the water–energy dynamics model and the refitted F&C's model (table 3; electronic supplementary material, figure S3 and table S3). The Moran's $I$ of the richness patterns of the four species groups was all dramatically reduced by our models
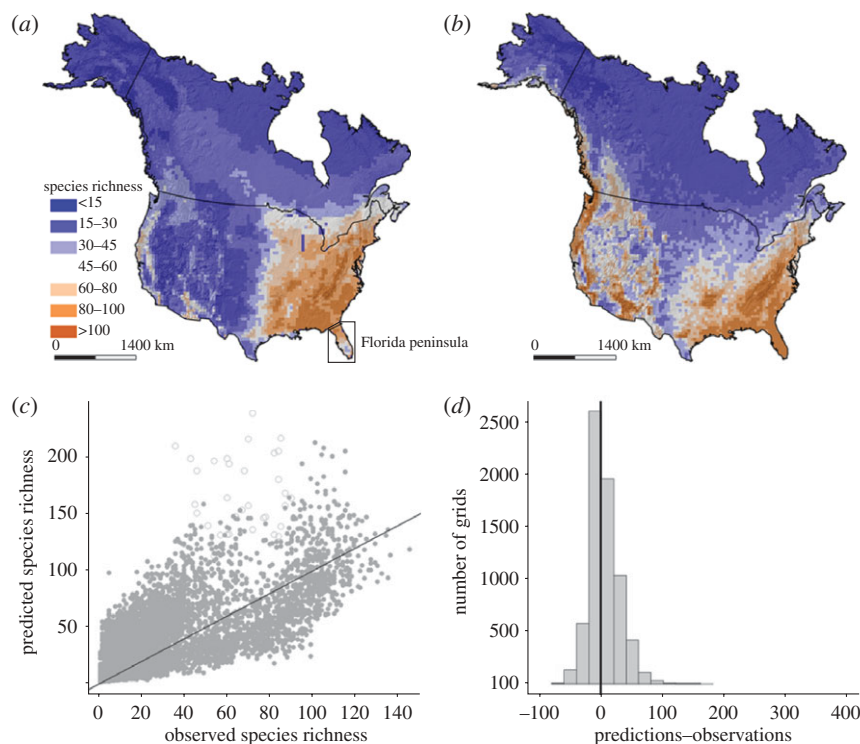
Figure 3. (*a,b*) Geographical patterns of the observed species richness of trees in North America (*a*) and the predictions of our model (*b*). (*c,d*) The relationships between the predictions and observations and the frequency distribution of the distances between the predictions and the 1 : 1 line. (*c*) Circles represent the grids in the Florida peninsula (demonstrated in (*a*)), and the solid line is the 1 : 1 line.

(electronic supplementary material, figure S1). In particular, the first-order Moran's $I$ (distance = 200 km) declined from 0.71–0.76 for species richness to 0.18–0.21 for model residuals. At larger distances, the residual Moran's $I$ was negligible and consistently lower than those of the other two models. Therefore, our model accounted for the major spatial structures in richness patterns.

Our model successfully predicted the species richness patterns of trees in North America (figure 3). The predictions were 36 (range: 1–238) per grid averagely, which was very close to the observations (average: 32; range: 1–145) [12,28,37]. The frequency distribution of the distances between the predictions and 1 : 1 line was symmetric (skewness = 0.88) and centred at five (s.d. = 24.2; figure 3), suggesting an average error of five species in the predictions. A special case is the Florida peninsula, where our model greatly overestimated the species richness (figure 3c). In contrast to our model, the water–energy dynamics model and the refitted F&C's model strongly overestimated the tree species richness in North America (electronic supplementary material, figure S4): the averages of their predictions were 52 (range: 0–1334) and 43 (range: 2–371) per grid, respectively. The frequency distributions of the distances between the predictions of these two models and 1 : 1 lines were highly right skewed (skewness = 9.83 and 2.74, respectively), with the averages being 20 (s.d. = 43.3) and 38 (s.d. = 46.9), respectively.

In the continental North Hemisphere, the species richness patterns of the four species groups predicted by our model were all consistent with the observed or predicted richness patterns of trees and vascular plants in previous studies [8,37,42]: species richness decreased from the equator to the North Pole, and was the highest in Central America, Southeast Asia and tropical Africa, but the lowest in Sahara, central Asia and boreal regions (figure 4).

## 4. DISCUSSION

### (a) *Comparison among climate, habitat heterogeneity and human activities*

Strong correlations between species richness and climate have been widely observed [7–9,13,28]. In our analyses, partial regressions indicate that the independent explanatory power of habitat heterogeneity and human activities is only *ca* 1/10 of those of climate, suggesting that climate is potentially the primary driver of large-scale patterns of species richness. In particular, our analysis showed that less than 4 per cent of the variation in the species richness of woody plants could be accounted for by individual variables of human activities, and less than 5–7% by the multiple models with all the human-activity variables as predictors. Similar results have been observed in Canada where the butterfly species richness is weakly correlated with human activities [27]. Such a low explanatory power of human activities may be owing to the influence of spatial scales. It has been suggested that the effects of human activities on species richness decrease with the expansion of study area [43].

Although the multiple regressions involving all the variables of habitat heterogeneity provided considerable explanatory power (46–54%) for the spatial variation in woody species richness, the two aspects of habitat heterogeneity, i.e. topographical relief and climatic heterogeneity, contributed differently [25]. Partial
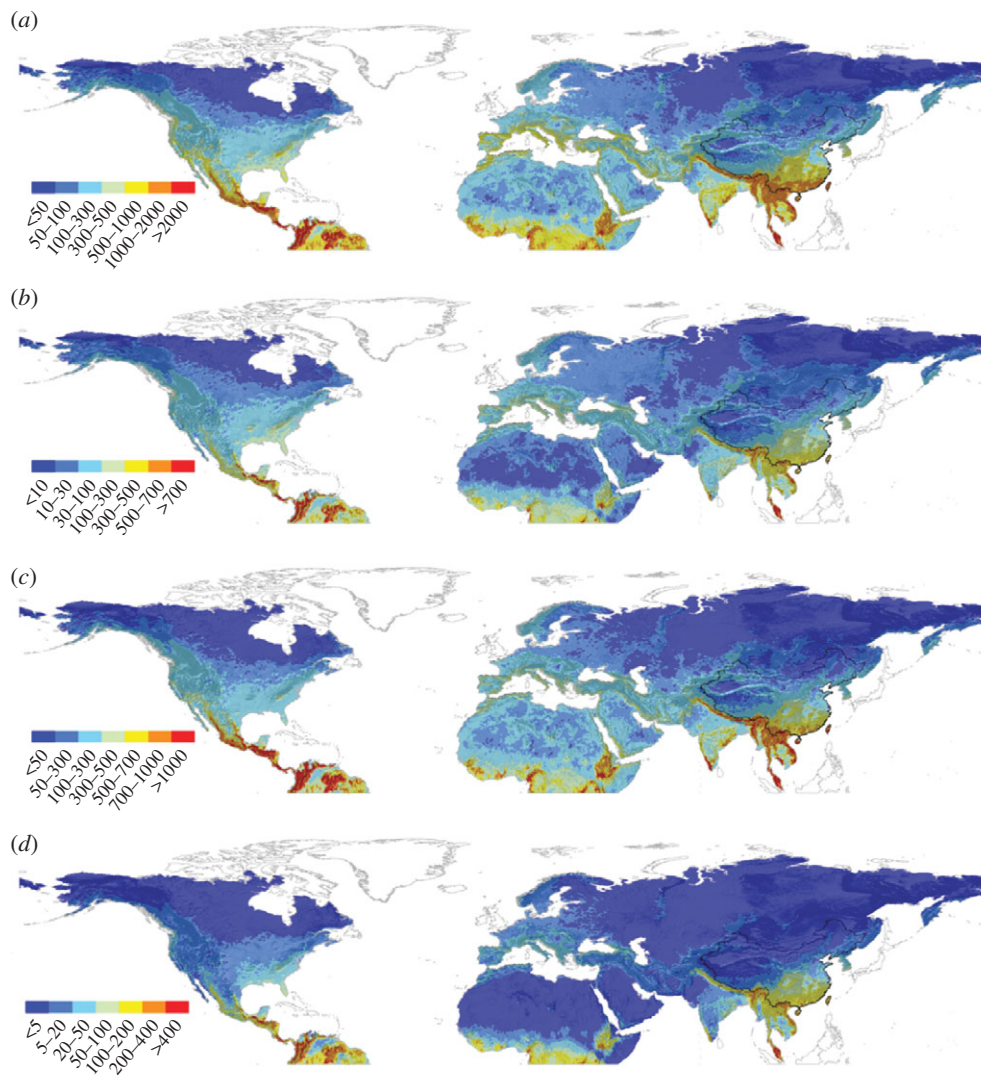
Figure 4. Species richness patterns of (*a*) all woody plants, (*b*) trees, (*c*) shrubs and (*d*) woody lianas in the Northern Hemisphere predicted by the model developed here.

regressions indicate that local climatic heterogeneity (i.e. RMAP and RMAT) independently accounted for 31–35% of richness variation if the effects of elevational range were first controlled, whereas elevational range independently explained less than 1 per cent of richness variation when the effects of local climatic heterogeneity were first controlled. This suggests that local climatic heterogeneity has stronger explanatory power for species richness than topographic relief. Similarly, previous studies indicate that the bird richness patterns in the mountains of the western Americas primarily reflected the effects of altitudinal variation of climate [25].

### (b) *Freezing-tolerance hypothesis*

Previous studies have indicated that energy, overwhelming water and climatic variability, is the most important climatic factor in determining large-scale patterns of species richness [2,11–14]. The hypotheses proposed to explain the mechanisms underlying the energy effects focus on different aspects of environmental energy, including ambient energy, chemical energy and winter coldness [2,11,20,21]. For example, Currie [2] found that ambient energy, instead of chemical energy, was the primary determinant for the diversity patterns of

vertebrates in North America, hence supporting the ambient energy hypothesis. However, he did not incorporate the variables of winter coldness. Hawkins *et al.* [13] compared the effects of ambient energy, chemical energy and winter coldness on the global pattern of bird diversity, and found that chemical energy had stronger effects than the other two, which supported the species richness-productivity hypothesis.

In contrast to previous studies, our analyses indicate that winter coldness (represented by MTCQ) accounted for much more variation in species richness than the other individual energy variables. In addition, the effects of winter coldness are stronger for species with tropical affinities than those with temperate affinities (table 2). More interestingly, although the two water–energy dynamics functions both have three variables, their effects on species richness are significantly lower than those of MTCQ for most species groups. One of the energy variables used in these combined functions, $PET_{min}$, is strongly correlated with MTCQ in the regions where both variables are above zero (the Northern Hemisphere: $r = 0.85$, $p < 0.001$; China: $r = 0.93$, $p < 0.001$), but is constantly zero in the regions where MTCQ is below zero (electronic supplementary material, figure S5). Specifically, $PET_{min}$ is zero in 73 per cent of the terrestrial
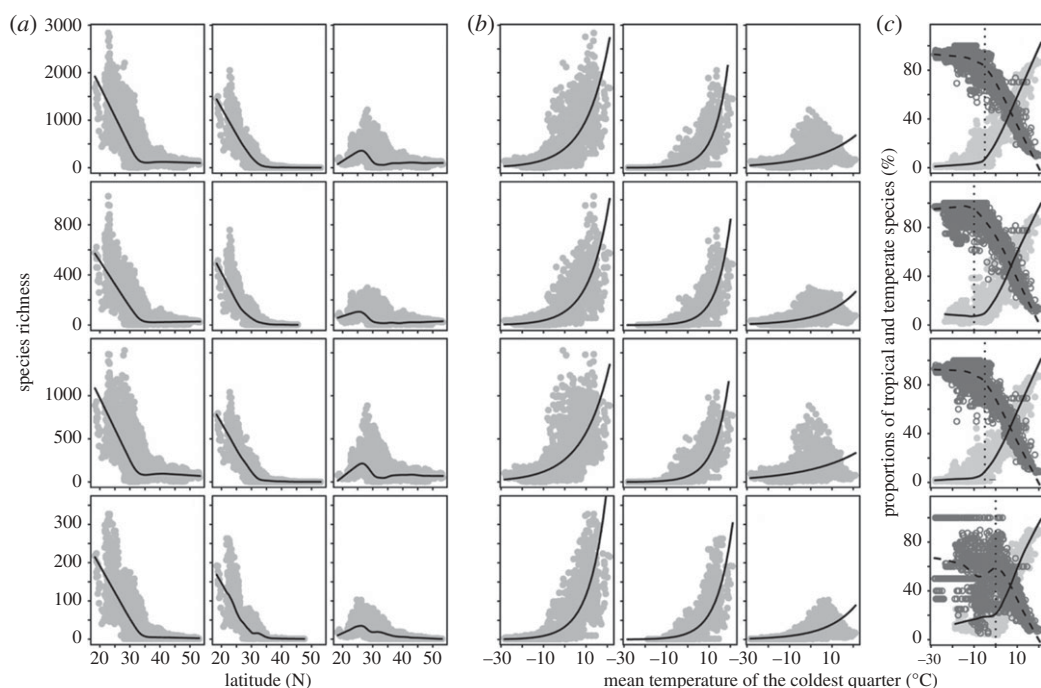
Figure 5. (*a,b*) The richness and (*c*) proportions of the species with tropical (dots and solid lines in (*c*)) and temperate (circles and dashed lines in (*c*)) affinities along the gradients of latitude and the mean temperature of the coldest quarter (MTCQ). In (*a,b*) the first column of graphs represents species richness; the second column, tropical; and the last column, temperate. From the top down, the four rows represent all woody plants, trees, shrubs and woody lianas, respectively. The solid lines in (*a,c*) are predicted by LOWESS method, and those in (*b*) by generalized linear models with Poisson residuals.

lands in China and 82.4 per cent in North America. Additionally, the two combined functions both strongly suffer from multi-collinearity (electronic supplementary material, appendix S1 and table S2) caused by the strong correlations between water and energy variables. These results may have reduced the performance of the water–energy dynamics functions in explaining species richness patterns [44]. Therefore, our results strongly support the recently formalized freezing-tolerance hypothesis (or tropical conservatism hypothesis) [20–22].

The theoretical framework of this hypothesis combines the effects of winter coldness with the niche conservatism of species, and emphasizes the difference in the tolerance of species which have evolved in different climates [20–22]. Our results indicate that the overall species richness is more strongly associated with that of tropical than temperate species for all woody plants and different life-forms, which suggests that the latitudinal gradient of species richness is mainly the results of the rapid decrease in the species richness with tropical affinities. This finding was confirmed by further comparisons between the patterns of overall, tropical and temperate species richness (figure 5). For all woody plants, with the increase of latitude and decrease of MTCQ, the richness of species with tropical affinity decreases much faster than that of the temperate-affinity species. Specifically, the overall and tropical species richness both dramatically decrease towards the north, while the richness of temperate species is the highest at latitudes ranging from 25° N to 30° N, and decreases towards the south and north. Additionally, the proportion of the species with tropical affinity in all species rapidly declines with the decrease of MTCQ, while the proportion of the species with temperate affinity rapidly increases (figure 5*c*). Generally, there are more

tropical than temperate species in grids with MTCQ of >5°C, but more temperate ones in grids with MTCQ of <5°C. Similar patterns are also observed for trees, shrubs and lianas (figure 5).

The opposite trends in the proportions of tropical and temperate species reflect their different tolerance of frost [20–22]. Most clades with tropical affinities have no traits of frost tolerance, and hence can be rapidly filtered out from local floras by the increasing winter coldness towards the poles. By contrast, the temperate clades are much less sensitive to winter coldness. For example, previous studies have indicated that evergreen broad-leaved trees in tropical and subtropical regions are very sensitive to winter temperature. Most evergreen broad-leaved trees in tropical rain forests cannot survive a temperature of <0°C, and those in subtropical forests would die at a temperature of < − 10°C [45,46]. By contrast, most of the tree species in temperate and boreal forests can tolerate the low temperature of −45°C to −60°C [45,46].

In summary, our results suggest that the changes in species richness are mainly the result of winter-coldness filtering for the species with tropical affinities that evolved in an ancient tropical-like climate and are sensitive to frost [22]. The control of winter coldness on species and vegetation distributions has long been observed in eastern Asia [36,45].

(c) *Models for predicting woody plant richness*
Our model for the species richness of woody plants has the same number of variables compared with the previously developed global model, IGM2 [17], and has one more variable than F&C's model [16]. However, our model has much higher explanatory power than the

other two models refitted in terms of China. In our model, MTCQ is the most important variable, which represents the winter coldness and individually accounts for 60–73% of richness variation (table 2). The other three variables, i.e. WD, temperature seasonality and the RMAT, improve the model's $r^2$ by 10–20%. The effects of winter coldness root in the evolution of species [20–22]; therefore, our model combines the effects of contemporary climate with the long-term evolutionary history. This may explain why it, using the coefficients developed in China, can successfully predict the species richness of trees in continental North America and woody plants in the North Hemisphere in spite of the huge difference in evolutionary history between different continents. WD was selected to represent the effects of water availability on species richness, which is consistent with F&C's model for the global pattern of angiosperm family richness [16]. However, the spatial correlograms showed that some spatial structures at small scales (distance < 200 km) can still be observed in the residuals of the model (electronic supplementary material, figure S1). These spatial structures may reflect the effects of other factors (e.g. soil), which are not included in this analysis. The overestimation of species richness in the Florida peninsula (figure 3c) may also reflect the effects of other factors. The low species richness of different taxa in the Florida peninsula has been observed for several decades, which may be caused by the peninsula effects [47]. Moreover, in the future, testing the model performance in the South Hemisphere will be helpful.

## REFERENCES

1 Rosenzweig, M. L. 1995 *Species diversity in space and time.* Cambridge, UK: Cambridge University Press.

2 Currie, D. J. 1991 Energy and large-scale patterns of animal- and plant-species richness. *Am. Nat.* **137**, 27–49. (doi:10.1086/285144)

3 Schipper, J. *et al.* 2008 The status of the world's land and marine mammals: diversity, threat, and knowledge. *Science* **322**, 225–230. (doi:10.1126/science.1165115)

4 Rahbek, C. & Graves, G. R. 2001 Multiscale assessment of patterns of avian species richness. *Proc. Natl Acad. Sci. USA* **98**, 4534–4539. (doi:10.1073/pnas.071034898)

5 Jetz, W. & Rahbek, C. 2002 Geographic range size and determinants of Avian species richness. *Science* **297**, 1548–1551. (doi:10.1126/science.1072779)

6 Orme, C. D. L. *et al.* 2005 Global hotspots of species richness are not congruent with endemism or threat. *Nature* **436**, 1016–1019. (doi:10.1038/nature03850)

7 Buckley, L. B. & Jetz, W. 2007 Environmental and historical constraints on global patterns of amphibian richness. *Proc. R. Soc. B* **274**, 1167–1173. (doi:10.1098/rspb.2006.0436)

8 Kreft, H. & Jetz, W. 2007 Global patterns and determinants of vascular plant diversity. *Proc. Natl Acad. Sci. USA* **104**, 5925–5930. (doi:10.1073/pnas.0608361104)

9 Currie, D. J. *et al.* 2004 Predictions and tests of climate-based hypotheses of broad-scale variation in taxonomic richness. *Ecol. Lett.* **7**, 1121–1134. (doi:10.1111/j.1461-0248.2004.00671.x)

10 Clarke, A. & Gaston, K. J. 2006 Climate, energy and diversity. *Proc. R. Soc. B* **273**, 2257–2266. (doi:10.1098/rspb.2006.3545)

11 Wright, D. H. 1983 Species-energy theory: an extension of species-area theory. *Oikos* **41**, 496–506. (doi:10.2307/3544109)

12 Currie, D. J. & Paquin, V. 1987 Large-scale biogeographical patterns of species richness of trees. *Nature* **329**, 326–327. (doi:10.1038/329326a0)

13 Hawkins, B. A., Porter, E. E. & Diniz-Filho, J. A. F. 2003 Productivity and history as predictors of the latitudinal diversity gradient of terrestrial birds. *Ecology* **84**, 1608–1623. (doi:10.1890/0012-9658(2003)084[1608:PAHAPO]2.0.CO;2)

14 Turner, J. R. G., Gatehouse, C. M. & Corey, C. A. 1987 Does solar energy control organic diversity? Butterflies, moths and the British climate. *Oikos* **48**, 195–205. (doi:10.2307/3565855)

15 O'Brien, E. M. 1998 Water–energy dynamics, climate, and prediction of woody plant species richness: an interim general model. *J. Biogeogr.* **25**, 379–398. (doi:10.1046/j.1365-2699.1998.252166.x)

16 Francis, A. P. & Currie, D. J. 2003 A globally consistent richness-climate relationship for angiosperms. *Am. Nat.* **161**, 523–536. (doi:10.1086/368223)

17 Field, R., O'Brien, E. M. & Whittaker, R. J. 2005 Global models for predicting woody plant richness from climate: development and evaluation. *Ecology* **86**, 2263–2277. (doi:10.1890/04-1910)

18 Connell, J. H. & Orias, E. 1964 The ecological regulation of species diversity. *Am. Nat.* **98**, 399–414. (doi:10.1086/282335)

19 O'Brien, E. M. 1993 Climatic gradients in woody plant species richness: towards an explanation based on an analysis of Southern Africa's woody flora. *J. Biogeogr.* **20**, 181–198. (doi:10.2307/2845670)

20 Latham, R. E. & Ricklefs, R. E. 1993 Global patterns of tree species richness in moist forests: energy-diversity theory does not account for variation in species richness. *Oikos* **67**, 325–333. (doi:10.2307/3545479)

21 Latham, R. & Ricklefs, R. 1993 Continental comparisons of temperate-zone tree species diversity. In *Species diversity in ecological communities: historical and geographical perspectives* (eds R. E. Ricklefs & D. Schluter), pp. 294–314. Chicago, IL: University of Chicago Press.

22 Wiens, J. J. & Donoghue, M. J. 2004 Historical biogeography, ecology, and species richness. *Trends Ecol. Evol.* **19**, 639–644. (doi:10.1016/j.tree.2004.09.011)

23 Kerr, J. T. & Packer, L. 1997 Habitat heterogeneity as a determinant of mammal species richness in high-energy regions. *Nature* **385**, 252–254. (doi:10.1038/385252a0)

24 Qian, H. & Ricklefs, R. E. 2000 Large-scale processes and the Asian bias in species diversity of temperate plants. *Nature* **407**, 180–182. (doi:10.1038/35025052)

25 Ruggiero, A. & Hawkins, B. A. 2008 Why do mountains support so many species of birds? *Ecography* **31**, 306–315. (doi:10.1111/j.2008.0906-7590.05333.x)

26 Fang, J. *et al.* 2006 Biodiversity changes in the lakes of the Central Yangtze. *Front. Ecol. Environ.* **4**, 369–377. (doi:10.1890/1540-9295(2006)004[0369:BCITLO]2.0.CO;2)

27 White, P. J. T. & Kerr, J. T. 2007 Human impacts on environment-diversity relationships: evidence for biotic homogenization from butterfly species richness patterns. *Global Ecol. Biogeogr.* **16**, 290–299. (doi:10.1111/j.1466-8238.2007.00298.x)

28 Wang, Z., Brown, J. H., Tang, Z. & Fang, J. 2009 Temperature dependence, spatial scale, and tree species diversity in eastern Asia and North America. *Proc. Natl Acad. Sci. USA* **106**, 13 388–13 392. (doi:10.1073/pnas.0905030106)

29 Fang, J., Wang, Z. & Tang, Z. 2010 *Atlas of woody plants in China: distribution and climate*. Berlin, Germany: Springer.

30 Fu, L. G. (ed.) 1999–2005 *Higher plants of China*. Qingdao, Shandong, China: Qingdao Publishing House.

31 Wu, Z. Y., Zhou, Z. K., Sun, H., Li, D. Z. & Peng, H. 2006 *The areal-types of seed plants and their origin and differentiation*. Kunming, China: Yunnan Science & Technology Press.

32 Qian, H., Song, J.-S., Krestov, P., Guo, Q., Wu, Z., Shen, X. & Guo, X. 2003 Large-scale phytogeographical patterns in East Asia in relation to latitudinal and climatic gradients. *J. Biogeogr.* **30**, 129–141. (doi:10.1046/j.1365-2699.2003.00807.x)

33 Harrison, S. & Grace, J. 2007 Biogeographic affinity helps explain productivity-richness relationships at regional and local scales. *Am. Nat.* **170**, S5–S15. (doi:10.1086/519010)

34 Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G. & Jarvis, A. 2005 Very high resolution interpolated climate surfaces for global land areas. *Int. J. Climatol.* **25**, 1965–1978. (doi:10.1002/joc.1276)

35 Thornthwaite, C. W. & Hare, F. K. 1955 Climatic classification in forest. *Unasylva* **9**, 51–59.

36 Kira, T. 1945 *A new classification of climate in Eastern Asia as the basis for agricultural geography*. Kyoto, Japan: Horticultural Institute Kyoto University.

37 Montoya, D., Rodríguez, M. A., Zavala, M. A. & Hawkins, B. A. 2007 Contemporary richness of holarctic trees and the historical pattern of glacial retreat. *Ecography* **30**, 173–182. (doi:10.1111/j.2006.0906-7590.04873.x)

38 Hou, Y. 2001 *Vegetation atlas of China*. Beijing, China: Science Press.

39 Dobson, A. 2002 *An introduction to generalized linear models*. London, UK: Chapman & Hall.

40 Legendre, P. & Legendre, L. 1998 *Numerical ecology*, 2nd English edn. Amsterdam, The Netherlands: Elsevier Science BV.

41 Fortin, M. & Dale, M. 2005 *Spatial analysis: a guide for ecologists*. Cambridge, UK: Cambridge University Press.

42 Barthlott, W., Lauer, W. & Placke, A. 1996 Global distribution of species diversity in vascular plants: towards a world map of phytodiversity. *Erdkunde* **50**, 317–327.

43 Sarr, D. A. 2005 A hierarchical perspective of plant diversity. *Q. Rev. Biol.* **80**, 187–212. (doi:10.1086/433058)

44 Hawkins, B. A., Montoya, D., Rodríguez, M. Á., Olalla-Tárraga, M. Á. & Zavala, M. Á. 2007 Global models for predicting woody plant richness from climate: comment. *Ecology* **88**, 255–259. (doi:10.1890/04-1910)

45 Sakai, A. 1979 Freezing tolerance of evergreen and deciduous broad-leaved trees in Japan with reference to tree regions. *Low Temp. Sci. Ser. B Biol. Sci.* **36**, 1–19.

46 Woodward, F. I. 1987 *Climate and plant distribution*. Cambridge, UK: Cambridge University Press.

47 Simpson, G. G. 1964 Species density of North American recent mammals. *Syst. Zool.* **13**, 57–73. (doi:10.2307/2411825)