

ENVIRONMENT/ECOLOGY

Analogous environments across the tropics have similar levels of tree species alpha diversity

Shumei Xiao ^{1,2,3}, Jonathan M. Adams ^{4,*}, Shufeng Li ^{1,2,100,123,*}, Ferry Slik ⁵, Daniel M. Griffith ⁶, Adriano Quaresma ⁷, Aisha Sultana ⁸, Andes Hamuraby Rozak ⁹, Andres Avella Muñoz ¹⁰, Andrew R. Marshall ^{11,12,13}, Arellano Gabriel ^{14,15}, Ashaq Ahmad Dar ¹⁶, Asyraf Mansor ^{17,18}, Ayyappan Narayanan ¹⁹, Bruno Herault ²⁰, Carlos Alfredo Joly ²¹, Daniel Piotto ²², David J. Harris ²³, Donald R. Drake ²⁴, Douglas Sheil ²⁵, Diogo S.B. Rocha ²⁶, Eddie Lenza de Oliveira ²⁷, Eddy Nurtjahya ^{28,29}, Eduardo van den Berg ³⁰, Edward L. Webb ^{31,32}, Faridah Hanum Ibrahim ³³, Felipe Zamborlini Saiter ^{34,35}, Florian Wittmann ³⁶, Francisco Mora Ardila ³⁷, Giselda Durigan ³⁸, Gopal Shukla ³⁹, Guillermo Ibarra-Manríquez ³⁷, Hidetoshi Nagamasu ⁴⁰, Ida Theilade ⁴¹, Irie Casimir Zo-Bi ⁴², Isau Huamantupa-Chuquimaco ^{43,44}, J. Orlando Rangel-Ch. ⁴⁵, James Grogan ⁴⁶, Javid Ahmad Dar ^{47,48}, Jochen Schöngart ⁴⁹, John Herbohn ⁵⁰, John R. Poulsen ⁵¹, John N. Williams ⁵², Jon Lovett ⁵³, Jose Don De Alban ^{54,55}, José Rafael Lozada ⁵⁶, José Roberto Rodrigues Pinto ⁵⁷, Juan Ernesto Guevara-Andino ⁵⁸, Jurgi Cristóbal-Azkarate ⁵⁹, Jürgen Homeier ⁶⁰, Katrin Böhning-Gaese ⁶¹, Khalid Rehman Hakeem ^{62,63,64}, Kenneth Feeley ⁶⁵, Kyle W. Tomlinson ⁶⁶, Ladan Rasingam ⁶⁷, Layon Oreste Demarchi ⁴⁹, Yves Laumonier ⁶⁸, Luciana F. Alves ⁶⁹, Luis Torres Montenegro ⁷⁰, Manichanh Satdichanh ⁷¹, Manuel J. Macía ^{72,73}, Marcelo Tabarelli ⁷⁴, Marcio Seiji Suganuma ⁷⁵, Marcos Antonio Ríos Paredes ^{70,76}, Maria Teresa Fernandez Piedade ⁴⁹, Mark Schulze ⁷⁷, Mattheas van de Bult ⁷⁸, Meredith L. Bastian ^{79,80}, Mohammed Latif Khan ⁸¹, Mohammad Shah Hussain ⁸, Michael Kessler ⁸², Michael J. Lawes ⁸³, Miguel A. Munguía-Rosas ⁸⁴, Narayanaswamy Parthasarathy ¹⁶, Naret Seuaturian ⁸⁵, Naveen Babu Kanda ^{16,19,86}, Naveenkumar Jagadeesan ³¹, Nigel Pitman ⁸⁷, Onrizal Onrizal ⁸⁸, Ole R. Vetaas ⁸⁹, Pantaleo Munishi ⁹⁰, Phourin Chhang ⁹¹, Polyanna da Conceição Bispo ⁹², Rahmad Zakaria ⁹³, Rahayu Sukmaria Sukri ⁹⁴, Rajkumar Muthu ⁹⁵, Rama Chandra Prasad ⁹⁶, Ramachandran V.S. ⁹⁷, Rhett D. Harrison ⁹⁸, Rizza Karen Veridiano ^{99,55}, Robert Steinmetz ⁸⁵, Robin L. Chazdon ¹⁰¹, Roven Tumaneng ^{55,102}, Samir Gonçalves Rolim ²², S.C. Garkoti ¹⁰³, Selene Báez ¹⁰⁴, Serge Wich ¹⁰⁵, Sharif A. Mukul ^{106,107}, Shijo Joseph ¹⁰⁸, Simone Aparecida Vieira ¹⁰⁹, S. Muthuramkumar ¹¹⁰, Somaiah Sundarapandian ¹⁶, Sumit Chakravarty ¹¹¹, Subashree Kothandaraman ^{47,48}, Tânia Wendt ¹¹², Thiago Metzker ^{113,114}, Timothy Whitfeld ¹¹⁵, Tao Su ^{12,116}, Tinde R Van Andel ¹¹⁷, Verbeeck Hans ¹¹⁸, Víctor Arroyo-Rodríguez ^{37,119}, Wanlop Chutipong ¹²⁰, William F. Laurance ¹²¹, Yrma Andreina Carrero ¹²² and Zhekun Zhou ^{1,2}

¹Laboratory of Tropical Forest Ecology, Xishuangbanna Tropical Botanical Garden, Chinese Academy of Sciences, Mengla 666303, China;

²Yunnan Key Laboratory of Forest Ecosystem Stability and Global Change, Xishuangbanna Tropical Botanical Garden, Chinese Academy of Sciences, Mengla 666303, China;

³University of Chinese Academy of Sciences, Beijing 100049, China;

⁴School of Geography and Ocean Science, Nanjing University, Nanjing 210023, China;

⁵Environmental and Life Sciences, Faculty of Science, University Brunei Darussalam, Gadong BE1410, Brunei Darussalam;

⁶Departamento de Ciencias Biológicas y Agropecuarias, EcoSs Lab, Universidad Técnica Particular de Loja, Loja 110150, Ecuador;

⁷Grupo de Pesquisa Ecologia, Monitoramento e uso Sustentável de Áreas Úmidas (MAUA), Instituto Nacional de Pesquisas da Amazônia (INPA), Manaus 69067-375, Brazil;

⁸Biodiversity Parks Programme, CEMDE, Department of Environmental Studies, University of Delhi, Delhi 110007, India;

⁹Research Organization for Life Sciences and Environment, National Research and Innovation Agency (BRIN), Bogor 16911, Indonesia;

¹⁰Facultad del Medio Ambiente y Recursos Naturales, Universidad Distrital Francisco José de Caldas, Bogotá 110231, Colombia;

¹¹Forest Research Institute, University of the Sunshine Coast, Sippy Downs, QLD 4556, Australia;

¹²Reforest Africa, Mang'ula, Kilombero, Tanzania;

(Continued on next page)

¹³Forest Research Institute, University of the Sunshine Coast, Sippy Downs, QLD 4556, Australia;

¹⁴Reforest Africa, Mang'ula, Kilombero, Tanzania;

¹⁵Forest Research Institute, University of the Sunshine Coast, Sippy Downs, QLD 4556, Australia;

¹⁶Reforest Africa, Mang'ula, Kilombero, Tanzania;

¹⁷Forest Research Institute, University of the Sunshine Coast, Sippy Downs, QLD 4556, Australia;

¹⁸Reforest Africa, Mang'ula, Kilombero, Tanzania;

¹⁹Forest Research Institute, University of the Sunshine Coast, Sippy Downs, QLD 4556, Australia;

²⁰Reforest Africa, Mang'ula, Kilombero, Tanzania;

²¹Forest Research Institute, University of the Sunshine Coast, Sippy Downs, QLD 4556, Australia;

²²Reforest Africa, Mang'ula, Kilombero, Tanzania;

²³Forest Research Institute, University of the Sunshine Coast, Sippy Downs, QLD 4556, Australia;

²⁴Reforest Africa, Mang'ula, Kilombero, Tanzania;

²⁵Forest Research Institute, University of the Sunshine Coast, Sippy Downs, QLD 4556, Australia;

ABSTRACT

Different regions of the tropics vary in overall tree species diversity, with the tropical Americas exhibiting strikingly higher regional tree species richness than Africa and Southeast Asia. We investigated whether these differences also occur at the local scale and whether the environmental conditions associated with tree species richness are consistent across tropical regions despite highly dissimilar species pools. A spatial random forest model was trained by using a network of 429 1-hectare plots across the tropics, together with 24 environmental variables, to predict plot-level tree α diversity. A combination of climatic, soil and topographical variables explained $\sim 86\%$ of the variation in richness. Despite differences in regional species pools and the potentially disruptive effects of different geological, climatic and evolutionary histories, the relationship between environmental variables and local-scale tree species richness is closely similar across different continents. Our findings imply a pervasive role of niche-based mechanisms in structuring local tree species richness, regardless of the regional species assemblages. This pantropical convergence in the richness–environment relationship poses a challenge for ecology to explain.

Keywords: rainforest, tree richness, modeling, climate, sample survey

INTRODUCTION

High levels of tree species richness and diversity in tropical forests have long fascinated biologists, representing an enduring challenge to ecological theory. While many mechanisms have been proposed to explain how such high levels of diversity have arisen and are maintained, substantial uncertainty persists [1–4]. To test emerging hypotheses about the mechanisms underlying tropical tree diversity, an empirical approach is required to identify robust predictors of species richness variation across broad scales [5]. An ideal method to explore and explain patterns of tropical tree diversity is to compare standardized inventory plots [1]. In a pioneering study, Gentry [1] compared a relatively limited number of plots by using least-squares regression to reveal a general relationship between primary productivity and species richness. Other studies of woody plant species richness, reviewed in [6], have concentrated on analysing single regions separately or are extratropical, preventing the comparison of tropical tree diversity relationships between regions. Ricklefs and He [7] compared 47 forest plots globally, fewer than half of which were from the tropics, finding that consistently warm and moist climates favored higher richness. However, they detected significant regional variation in extratropical regions and the limited number of plots prevented definitive comparison within the tropics.

Consideration of the large differences in geological and evolutionary history between the different regions and subregions of tropical forest led us to propose two main hypotheses: (i) local-scale tree α diversity across the global tropics can be consistently predicted by using contemporary environmental variables, such that analogous environmental conditions will support similar levels

of richness, irrespective of deep historical biogeographic divergences; (ii) due to localized effects of forest history and dispersal lag, and the complex ecology of forest communities, there will be differences in the best predictors of 1-ha richness at the local scale compared with at the broader scale.

While more recent studies have expanded upon this foundation, pantropical comparisons of the patterns and potential drivers of local tree diversity remain limited. Comparing 2046 tree plots across Amazonia, Ter Steege *et al.* [8] demonstrated that a combined influence of climate and soil factors explained local tree richness and community composition. In a comparison of forest plots from South America and Africa, Parmentier *et al.* [9] showed lower plot-level richness in Africa under similar warm and moist climatic conditions. In global surveys of forest plots of varying sizes, Keil and Chase [10] and Chu *et al.* [11] found that the drivers of the variation in diversity differ with the plot size and spatial distance between samples, which are important sampling issues in ecology generally [12]. However, no studies thus far have addressed local tree species richness by using a large number of plots of standardized size across the world's tropics, which is necessary for continental-scale comparisons. The increasing availability and integration of forest plot data now make such comparisons possible [13].

Two other important advances in environmental data science facilitate hypothesis testing about the patterns and processes driving tropical tree diversity. First, the availability of interpolated climate, soil and other environmental parameters has grown immensely in recent years [14–18]. Second, machine-learning models enable analysis of the simultaneous effects of multiple factors on community structure and diversity across large scales [19,20]. A particularly

(Continued from previous page)

¹³Flamingo Land Ltd, Kirby Misperton, North Yorkshire, YO17 6UX, UK;

¹⁴Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, MI 48109, USA;

¹⁵Oikobit LLC, Albuquerque, NM 87120, USA;

¹⁶Department of Ecology and Environmental Sciences, School of Life Sciences, Pondicherry University, Puducherry 605014, India; ¹⁷School of Biological Sciences (SBS), Universiti Sains Malaysia, Penang 11800, Malaysia;

¹⁸Centre for Marine and Coastal Studies (CEMACS), Universiti Sains Malaysia, Penang 11800, Malaysia; ¹⁹Department of Ecology, French Institute of Pondicherry, Puducherry 605001, India; ²⁰CIRAD, Campus de Baillarguet, Montpellier Cedex 5, 34398, France; ²¹Department of Plant Biology, Institute of Biology, University of Campinas, Campinas 13083-862, Brazil;

²²Laboratório de Dendrologia e Silvicultura Tropical, Centro de Formação em Ciências Agroflorestais, Universidade Federal do sul da Bahia, Itabuna 45613-204, Brazil; ²³Royal Botanic Garden Edinburgh, Edinburgh EH3 5LR, UK; ²⁴Harold L. Lyon Arboretum, University of Hawai'i at Mānoa, Honolulu, HI 96822, USA;

²⁵Forest Ecology and Forest Management Group, Wageningen University & Research, Wageningen 6700 AA, The Netherlands; ²⁶International Institute for Sustainability, Rio de Janeiro 22460-320, Brazil;

²⁷Programa de Pós-graduação em Ecologia e Conservação, Universidade do Estado de Mato Grosso, Nova Xavantina 78690-000, Brazil; ²⁸Department of Biology, Universitas Bangka Belitung, Kampus Terpadu UBB Balunijuk, Kec.

Merawang, Kab. Bangka, 33172, Indonesia; ²⁹Herbarium Bangka Belitungense, Kampus Terpadu UBB Balunijuk, Kec.

Merawang, Kab. Bangka, 33172, Indonesia;

³⁰Herbarium Bangka Belitungense, Kampus Terpadu UBB Balunijuk, Kec.

Merawang, Kab. Bangka, 33172, Indonesia;

³¹Herbarium Bangka Belitungense, Kampus Terpadu UBB Balunijuk, Kec.

Merawang, Kab. Bangka, 33172, Indonesia;

(Continued from previous page)

³⁰Departamento de Ecologia and Conservation, Universidade Federal de Lavras, Lavras, MG 37200-900, Brazil; ³¹Viikki Tropical Resources Institute, Department of Forest Sciences, University of Helsinki, Helsinki 00790, Finland; ³²Helsinki Institute of Sustainability Science (HELSUS), Helsinki 00100, Finland; ³³Institute of Tropical Forestry and Forest Products (INTROP), Universiti Putra Malaysia, Serdang, Selangor 43400, Malaysia; ³⁴Federal Institute of Espirito Santo, Cariacica Campus, Espirito Santo 29150-410, Brazil; ³⁵National Institute of the Atlantic Forest, Santa Teresa, Espirito Santo 29650-000, Brazil; ³⁶Department of Wetland Ecology, Karlsruhe 76131, Germany; ³⁷Instituto de Investigaciones en Ecosistemas y Sustentabilidad, Universidad Nacional Autónoma de México, Morelia 58190, Mexico; ³⁸Floresta Estadual de Assis, Instituto de Pesquisas Ambientais, Assis 19800-000, Brazil; ³⁹Department of Forestry, North Eastern Hill University, Tura Campus, Tura, Meghalaya 794002, India; ⁴⁰The Kyoto University Museum, Kyoto University, Kyoto 606-8501, Japan; ⁴¹Department of Food and Resource Economics, University of Copenhagen, DK-1958 Frederiksberg C, Denmark; ⁴²Institute National Polytechnique Félix Houphouët-Boigny, INP-HB, Yamoussoukro, BP 1093, Côte d'Ivoire; ⁴³Herbario Alwyn Gentry (HAG), Departamento Académico de Ciencias Básicas, Universidad Nacional Amazónica de Madre de Dios (UNAMAD), Madre de Dios 17001, Peru; ⁴⁴Centro Ecológico INKAMAZONIA, Valle de Kosñipata, vía Cusco-Parque Nacional del Manú, Cusco, Peru; ⁴⁵Instituto de Ciencias Naturales, Universidad Nacional de Colombia, Bogotá D.C. 111321, Colombia; ⁴⁶Botanic Garden of Smith College, Northampton, MA 01063, USA;

robust machine-learning method is random forest (RF) modeling, which, when combined with spatial regression, enables the analysis and prediction of spatially structured data in which observations exhibit autocorrelation [21,22]. These models extend the capabilities of traditional RF to handle spatial data more effectively, allowing spatial autocorrelation and other geographic phenomena to be incorporated into vegetation modeling.

Here, we used the Pantropical Forests Network (PFN) of tree inventory plots assembled by Slik *et al.* [23] in combination with publicly available environmental data surfaces [14–18] to determine which environmental variables explain plot-level tree α diversity across the global tropics. We applied a spatial RF and negative binomial generalized linear model to detect and compare the strength of empirical links between environmental variables and tree species richness. The number of species per unit area (e.g. Gentry [1])—the classic measure of forest tree richness—has its weaknesses given that stem density varies by both biogeographic region and along climate gradients. To compare species richness while accounting for variations in sampling effort and completeness, we employed rarefaction, which is a sample-coverage-based approach based on abundance data (e.g. [24,25]), together with Fisher's α [26], and the 'classical' method of the number of species recorded per plot, with the goal of improving understanding of the potential drivers and maintenance of tropical diversity. Specifically, we addressed three main questions: (i) Which environmental predictors explain most of the variation in local tree species richness across the tropics? (ii) How do the environmental factors associated with tree species richness differ between broad versus local scales of sample plot spacing? (iii) Given the environmental predictors, does tree species richness at the 1-ha scale converge between tropical regions to give a consistent pantropical pattern?

A main goal of this study is to determine whether local tree richness and the environmental factors associated with it differ amongst the world's major tropical regions. Such a global comparison is important to inquire whether community assemblage rules may be consistent across geographically disparate regions—in other words, evolutionarily conserved underlying environment–richness relationships across the tropical forest biome.

RESULTS

A principal component analysis (PCA) conducted on 24 environmental variables across 1-ha plots

in the global tropics revealed regional differences, with the first two axes accounting for 40.5% of the variance in the variables (Table S1 and Fig. 1). Overall, the three regions exhibited moderate separation along the first principal component (PC1), especially between the Americas and Asia, but were largely overlapping along the second principal component (PC2). PC1 was primarily associated with moisture-related variables, including precipitation seasonality (bio15), the precipitation of the driest month (bio14) and the annual range of monthly relative humidity (hurs_range), along with soil pH (phh2o), isothermality (bio03), net primary productivity (NPP) and annual range and mean of monthly surface downwelling shortwave flux in air (rsds_range and rsds_mean) (Fig. 1 and Table S2). PC2 predominantly comprised soil-related variables, including total nitrogen, volumetric water content (wv0010) and soil organic carbon content in the fine earth fraction (SOC), as well as topography-related variables, including tangential curvature (tcurv), ruggedness and vector ruggedness measure (vrm) (Fig. 1 and Table S2). The Americas exhibited greater variation in environmental conditions along PC1, characterized by higher precipitation during the driest month and lower precipitation seasonality compared with Asia and Africa (Fig. 1). Specifically, Asia demonstrated the highest precipitation during the driest month and lower precipitation seasonality, while Africa showed the lowest variation in these environmental conditions. Variation along PC2 was similar among the three regions (Fig. 1).

We obtained the species richness estimated from rarefaction based on sample coverage (Fig. 2) and Fisher's α from a standardized sample of stems. By analysing non-spatial and spatial RF based on the observed richness, Fisher's α and the richness from rarefaction data returned high 'out-of-bag' R^2 values (0.83–0.88), demonstrating a robust ability to predict 1-ha-scale tree species richness from the 24 environmental predictors across the global tropics (Fig. 3 and Table S1).

Training the RF for each of the three richness measures revealed that the predicted tree species richness at the 1-ha scale was highly heterogeneous across the tropics (Fig. 4a and Figs S6a and S7a), with the predicted richness highest in western South America, particularly the Andean–Amazon foothills and Colombian Chocó, the major islands of Southeast Asia, and New Guinea (Fig 4a). Low levels of predicted local species diversity were found across most of tropical Africa, eastern and southern Amazonia and continental Southeast Asia.

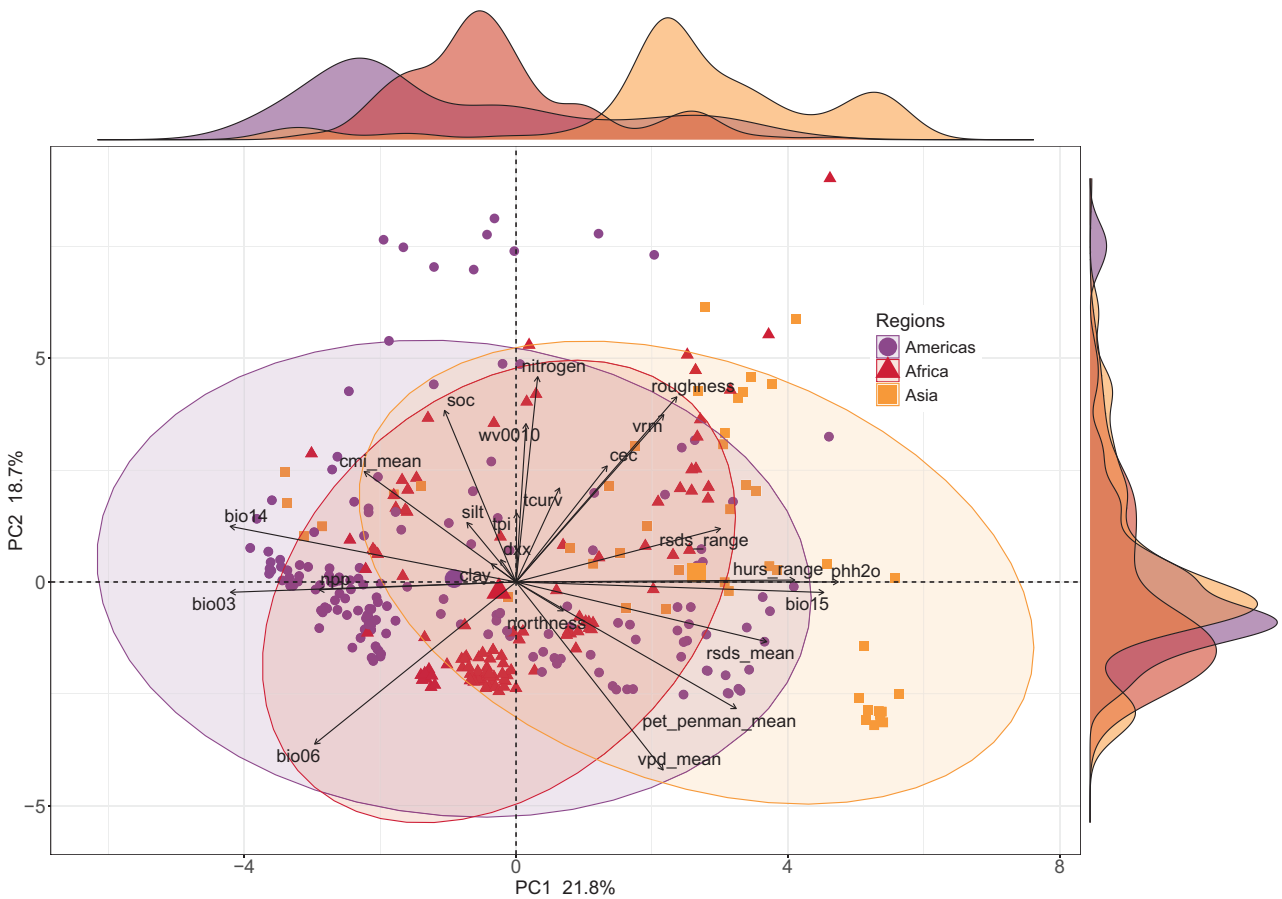


Figure 1. PCA of 24 environmental variables for 1-ha tropical tree plots in the Americas, Africa and Asia. Marginal density plots above and to the right of the biplot show the distribution of samples from each region along the first (PC1) and second (PC2) principal components, respectively. Percentages on each axis represent the variation explained by the respective principal component.

(Continued from previous page)

⁴⁷Department of Environmental Science and Engineering, School of Engineering and Sciences, SRM University-AP, Andhra Pradesh 522240, India; ⁴⁸Centre for Geospatial Technology, SRM University-AP, Andhra Pradesh 522240, India; ⁴⁹Ecology, Monitoring and Sustainable Use of Wetlands (MAUA), Instituto Nacional de Pesquisas da Amazônia (INPA), Manaus 69060-001, Brazil; ⁵⁰Forest Research Institute, University of the Sunshine Coast, Sippy Downs, Queensland 4556, Australia; ⁵¹The Nature Conservancy, Boulder, CO 80302, USA; ⁵²Department of Environmental Science and Policy, University of California, Davis, CA 95616, USA; ⁵³School of Geography, University of Leeds, Leeds LS2 9JT, UK;

Analysis using non-spatial RF revealed that bio14 was identified as the most important single variable for predicting tree richness, followed by phh2o (Fig. 4b). The interaction between the precipitation of the driest month and the proportion of silt particles in the fine earth fraction (silt) also contributed significantly to the model (Fig. 4b). Spatial RF, which accounted for spatial autocorrelation (see [Methods S1](#)), showed a higher importance of moisture-related variables, including hurs_range, bio15, the mean monthly potential evapotranspiration (pet_penman_mean) and the mean monthly climate moisture index (cmf_mean) (Fig. 4c). In contrast, the importance of soil-related variables, such as pH and silt, remained relatively consistent, while the interaction between the precipitation of the driest month and silt (bio14.pca..silt) decreased in importance. An additional analysis in which non-spatial and spatial RF were applied to a thinned set of sample locations separated by at least 50, 100 and 200 km showed a diminished importance of soil-related

variables such as pH, SOC and silt (Fig. 4d and [Fig. S8b](#) and d). This reduction was particularly pronounced with thinning of 100 and 200 km, for which the importance of soil pH and silt decreased markedly. In contrast, variables associated with moisture (e.g. hurs_range, cmf_mean), growing season (e.g. NPP), solar radiation (e.g. rds_mean) and topography (e.g. topographic position index [tpi]) increased in importance.

According to spatial RF, the Americas, Africa and Asia exhibited similar relationships between tree species richness and six dominant environmental variables: bio14, phh2o, hurs_range, bio15, silt and pet_penman_mean ([Fig. S9](#)). In all three regions, the tree richness was highest in areas with abundant moisture in the driest month (bio14; [Fig. S9a](#)). Conversely, the richness was low in areas with high soil pH and silt fraction, except in Asia, where the richness increased with silt ([Fig. S9b](#) and e). High tree richness was also associated with a low annual range of monthly near-surface relative humidity, precipitation

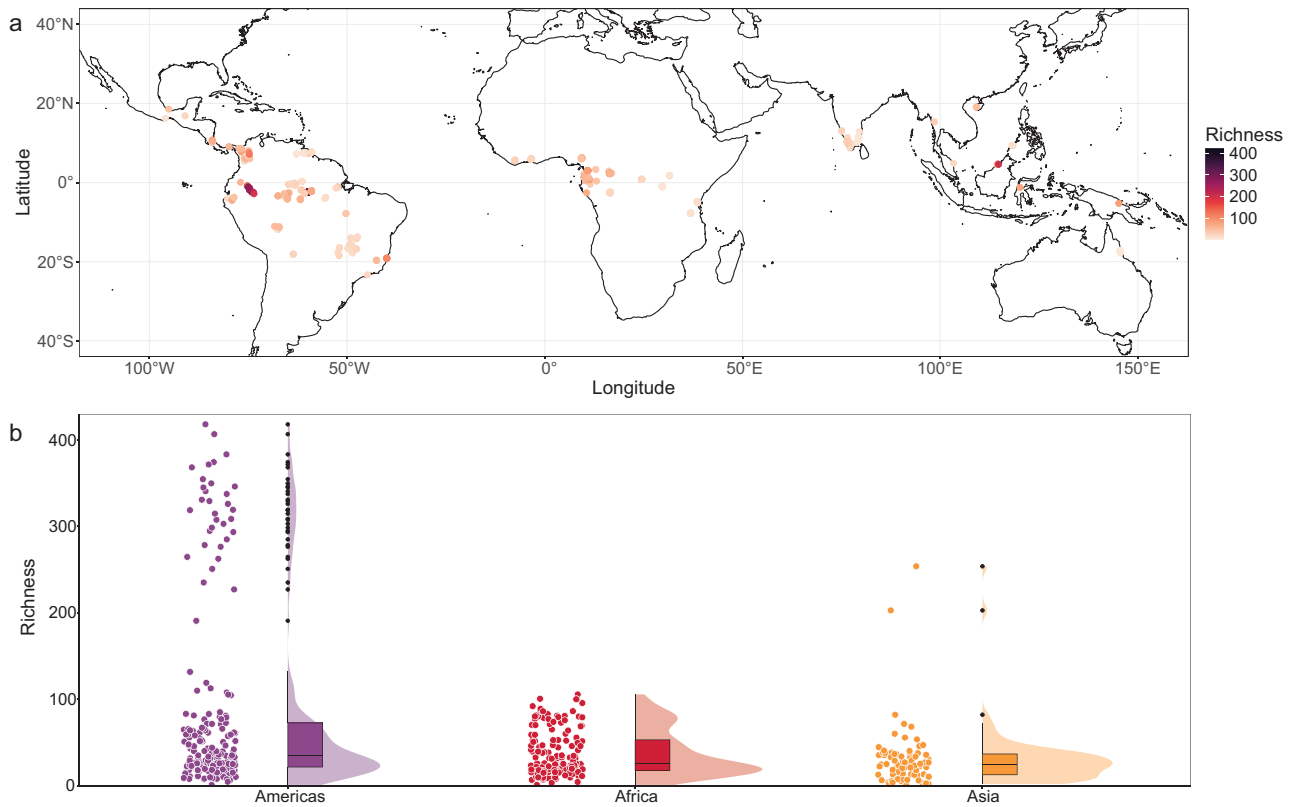


Figure 2. Tree species richness in 1-ha plots across the global tropics. (a) Distribution of 429 1-ha plots located in undisturbed old-growth forests used in this study. (b) Density distribution of the plots for the three regions. Box plots show the median and interquartile range of species richness estimated from rarefaction, alongside individual plot data for each region. Vertical lines extend to 1.5 times the difference between the quartiles and black points represent the outliers. Width of the distribution represents the number of plots at a given richness level. Similar plots for Fisher's α and observed richness are shown in Figs S3 and S4. Review drawing number: GS京(2025) 2372号.

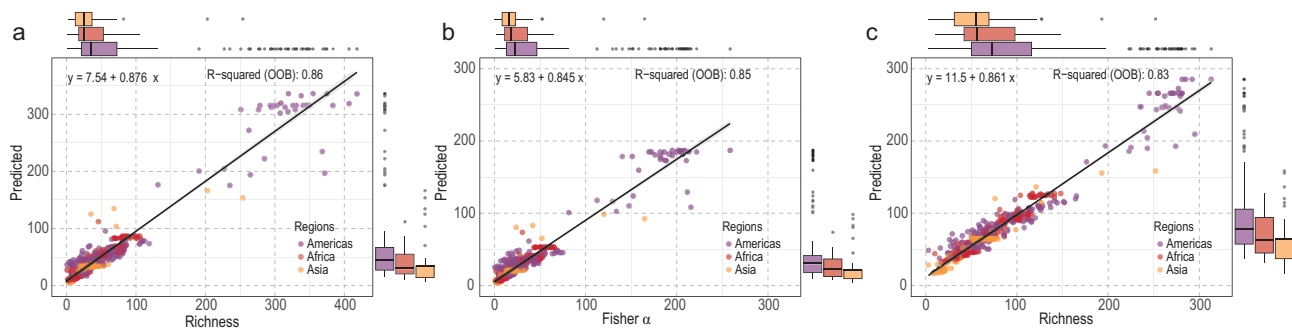


Figure 3. Predicted tree species richness according to non-spatial RF based on 24 environmental variables compared with richness estimated from (a) sample-based rarefaction, (b) Fisher's α and (c) observed richness. Tree inventory data were obtained from 429 1-ha plots of old-growth tropical forest located across the Americas, Africa and Asia. The out-of-bag R^2 reflects the performance of each RF model based on observations that were excluded from the training subset for each tree. Box plots above and to the right of each graph show the estimated and predicted richness between regions, respectively.

(Continued from previous page)

⁵⁴ Clear Wind Pte Ltd, Singapore 068894, Singapore; ⁵⁵ Fauna & Flora International—Philippines Programme, Cambridge CB2 3QZ, UK;

seasonality and mean monthly potential evapotranspiration, peaking when these were approximately 5%–6%, 25–35 mm and 110–115 mm/month, respectively (Fig. S9c, d and f). Notably, the Americas displayed a large variation in the precipitation of the driest month, soil pH and silt, whereas Asia exhibited substantial

variation in the humidity range, precipitation seasonality and monthly potential evapotranspiration. Africa showed a low variation in these parameters compared with the other two regions.

The contribution of the six dominant environmental variables to local tree species richness varied within and across the regions. In areas with

(Continued from previous page)

⁵⁶Instituto de Investigaciones para el Desarrollo Forestal (INDEFOR), Universidad de Los Andes, Mérida 5101, Venezuela; ⁵⁷Departamento de Engenharia Florestal, Centro de Referência em Conservação da Natureza e Recuperação de Áreas Degradadas, Universidade de Brasília, Brasília 70910-900, Brazil; ⁵⁸Grupo de Investigación en Ecología y Evolución en los Trópicos-EETrop, Universidad de las Américas, Quito 170125, Ecuador; ⁵⁹Department of Basic Psychological Processes and Development, University of the Basque Country UPV/EHU, Donostia 20018, Spain; ⁶⁰Faculty of Resource Management, HAWK University of Applied Sciences and Arts, Göttingen 37085, Germany; ⁶¹Senckenberg Biodiversity and Climate Research Centre, Frankfurt am Main, Germany and Goethe University, Frankfurt am Main 60323, Germany; ⁶²Department of Biological Sciences, Faculty of Science, King Abdulaziz University, Jeddah 21589, Saudi Arabia; ⁶³Princess Dr. Najla Bint Saud Al-Saud Center for Excellence Research in Biotechnology, King Abdulaziz University, Jeddah 21589, Saudi Arabia; ⁶⁴Centre of Research Impact and Outcome, Chitkara University Institute of Engineering and Technology, Chitkara University, Rajpura, Punjab 140401, India; ⁶⁵Biology Department, University of Miami, Coral Gables, FL 33146, USA; ⁶⁶Center for Integrative Conservation & Yunnan Key Laboratory for Conservation of Tropical Rainforests and Asian Elephants, Xishuangbanna Tropical Botanical Garden, Mengla 666303, China; ⁶⁷Botanical Survey of India, Deccan Regional Centre, Hyderabad, Telangana 500095, India; ⁶⁸Centre for International Forestry Research, Bogor 16115, Indonesia; ⁶⁹Institute of the Environment and Sustainability, University of California Los Angeles, Los Angeles, CA 90095, USA; ⁷⁰Herbario Herrerense, Instituto de Investigaciones de la Amazonia Peruana, Iquitos 16002, Peru;

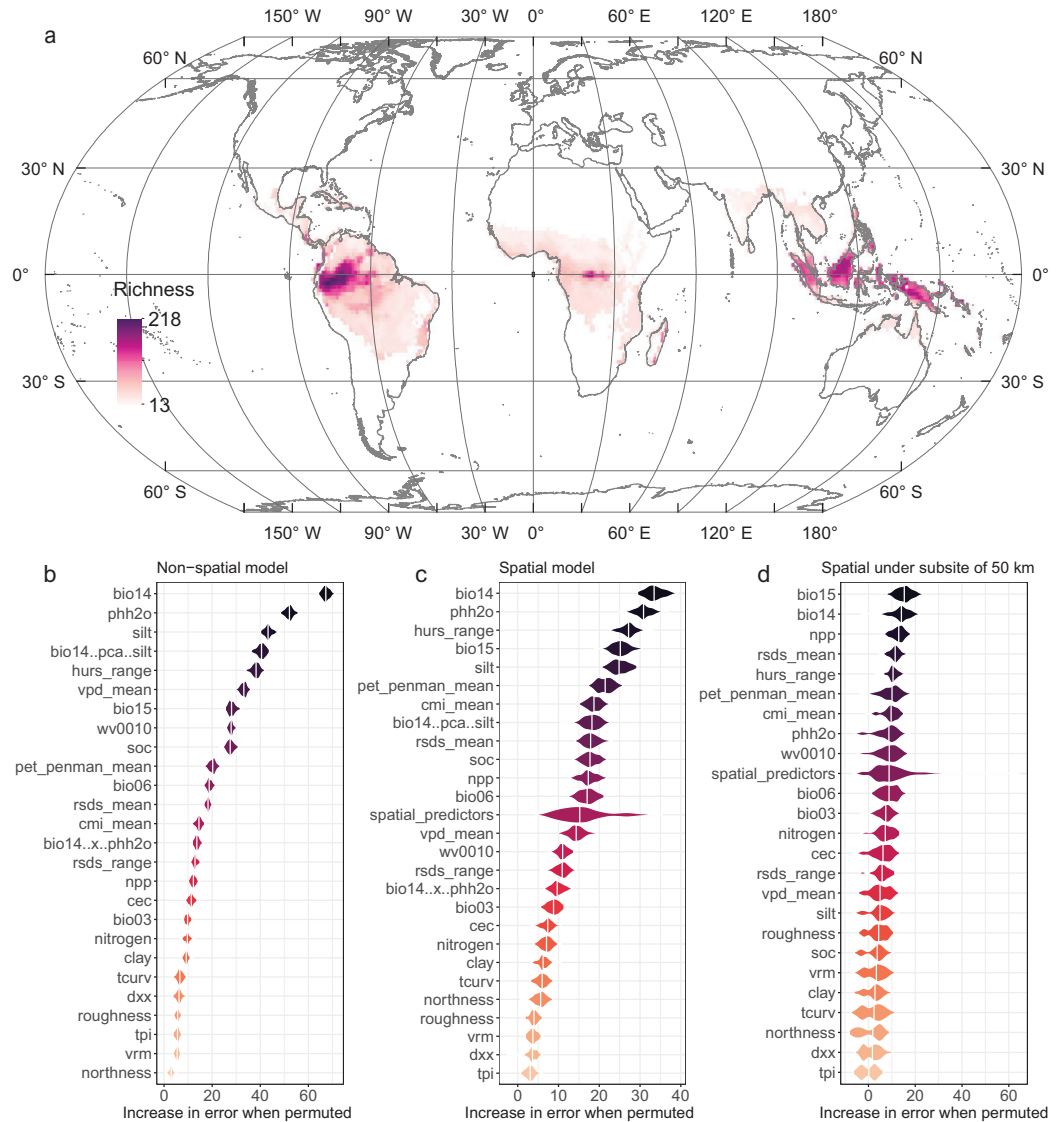


Figure 4. Predicted tree species richness at the 1-ha scale and importance of environmental variables based on non-spatial and spatial RF. (a) Local tree species richness estimated from rarefaction and predicted using non-spatial RF across the tropics. (b) Importance of environmental variables according to non-spatial RF, which does not account for spatial autocorrelation among forest plots. (c) Importance of environmental variables according to spatial RF, which accounts for spatial autocorrelation among forest plots. (d) Importance of environmental variables according to spatial RF with thinning of samples by 50 km. Review drawing number: GS京(2025) 2372号.

high precipitation, low seasonality in rainfall and low soil pH (Figs S10 and S11), the predictive contribution of these climatic variables to species richness within the RF model was reduced (Figs S12 and S13). For example, in areas with low soil pH (~3.7–4), such as the Amazon (Fig. S10b), the importance of soil pH in the spatial RF decreased (Fig. S12b). By contrast, in areas characterized by high soil pH, such as India and southeast Africa, this soil parameter was a critical predictor in the model (Fig. S12b).

The negative binomial generalized linear models (glm.nb) demonstrated a weaker ability

than the RF to predict local tree species richness, whether estimated from rarefaction (OOB $R^2 = 0.69$), Fisher's α (OOB $R^2 = 0.70$) or observed richness (OOB $R^2 = 0.72$) (Fig. S14). Moreover, in contrast to results from the spatial RF, the glm.nb found that Africa and Asia exhibited similar levels of local tree species richness, both of which were significantly lower than those of the Americas (Tables S6 and S7). Despite those differences, the direction and significance of the effects of important precipitation (e.g. bio14, bio15, mean monthly vapor pressure deficit [vpd_mean], hurs_range, temperature

(Continued from previous page)

⁷¹School of Life Sciences, University of Hawai'i at Mānoa, Honolulu, HI 96822, USA; ⁷²Departamento de Biología, Área de Botánica, Universidad Autónoma de Madrid, Madrid 28049, Spain; ⁷³Centro de Investigación en Biodiversidad y Cambio Global (CIBC-UAM), Universidad Autónoma de Madrid, Madrid 28049, Spain; ⁷⁴Botany Department, Universidade Federal de Pernambuco, Recife, PE 50670-901, Brazil; ⁷⁵Colégio de Aplicação, Universidade Federal de Santa Catarina, Florianópolis 88040-900, Brazil; ⁷⁶Keller Science Action Center, The Field Museum, Chicago, IL 60605-2496, USA; ⁷⁷H.J. Andrews Experimental Forest, Blue River, OR 97413, USA; ⁷⁸Doi Tung Development Project, Social Development Department, Chiang Rai 57240, Thailand; ⁷⁹Proceedings of the National Academy of Sciences, Washington, DC 20001, USA; ⁸⁰Department of Evolutionary Anthropology, Duke University, Durham, NC 27708, USA; ⁸¹Department of Botany, Dr. Harisingh Gour Vishwavidyalaya (A Central University), Sagar, Madhya Pradesh 470003, India; ⁸²Systematic and Evolutionary Botany, University of Zurich, Zurich 8008, Switzerland; ⁸³School of Life Sciences, University of KwaZulu-Natal, Scottsville 3209, South Africa; ⁸⁴Laboratorio de Ecología Terrestre, Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional (Cinvestav), Mérida 97310, Mexico; ⁸⁵World Wildlife Fund Thailand, Bangkok 10400, Thailand; ⁸⁶National Biobank of Thailand, National Science and Technology Development Agency, Pathum Thani 12120, Thailand; ⁸⁷Field Museum of Natural History, Chicago, IL 60605, USA; ⁸⁸Faculty of Forestry, Universitas Sumatera Utara, Medan 20155, Indonesia; ⁸⁹Department of Geography, University of Bergen, Bergen N-5020, Norway;

(mean monthly minimum air temperature of the coldest month [bio06]) and soil variables (e.g. pH, silt) were similar between the two approaches (Tables S6 and S7). Variables showing a significant interaction between Africa and the Americas included *cmi_mean*, *hurs_range*, *vpd_mean*, clay proportion, soil pH and *tcurv* (Tables S6 and S7). Both the *glm.nb* and RF models showed consistent results, indicating that variables such as *bio14*, pH, *hurs_range*, silt and *bio15* are important factors in controlling α diversity. While RF excels in predictive power ($R^2 = 0.86$), capturing complex, nonlinear relationships, the *glm.nb* model ($R^2 = 0.69$) also explains a large portion of the richness (Fig. 2 and Fig. S14). This demonstrates *glm.nb*'s ability to capture substantial data variability despite its lower R^2 . Crucially, *glm.nb* shows significant regional differences between the three regions, confirming broad-scale patterns. While the results from *glm.nb* and RF diverged in important ways, the RF model is given precedence here due to its superior OOB R^2 values and capacity to incorporate nonlinear effects [27,28]. However, the *glm.nb* model offers the advantage of interpretability and simplicity, which can be crucial for understanding the underlying factors driving richness. Therefore, while RF may provide superior predictive accuracy, *glm.nb* remains a valuable tool for future additional exploratory analysis and hypothesis testing.

When environmental variables were grouped into seven categories—temperature, precipitation, growing season, solar radiation, soil, topography and co-limitation, which means no single type of environmental factor dominates—the local tree species richness was largely explained by the same three categories: co-limitation, precipitation and soil (Fig. 5). Referring to areas in which no single category dominates, co-limitation represented the most important category, accounting for 30.02% of the variation in local tree species richness (Fig. 5b). Precipitation (22.90%) and soil (17.56%) were the next most important categories, followed by solar radiation (11.71%) and topography (11.10%). In contrast, temperature (2.05%) and growing season (4.65%) were relatively unimportant overall, but increased in significance at higher latitudes and altitudes within the tropics (Fig. 5a). Solar radiation and growing season tended to gain importance relative to other categories where the species richness was low, particularly near 10°N and 10–20°S latitude. Along the latitudinal gradient, co-limitation, moisture and soil were the dominant categories explaining the richness variation in the three regions. Compared with other regions, along the longitudinal gradient, the

growing season and solar radiation were more important in Africa, while the soil and precipitation categories were most important in Asia. Overall, species richness in most tropical regions is constrained by multiple environmental factors, indicating a co-limitation effect, particularly in areas of higher species richness close to the equator.

DISCUSSION

The relationship between environmental factors and species richness

The RF analysis and *glm.nb* model showed that variation in tropical tree species richness was consistently predictable by using 24 primary environmental variables across the three main biogeographic regions of the tropics. Local tree species richness predicted by RF were strongly correlated with those based on sample coverage estimation from rarefaction and Fisher's α across the tropics, with the observed raw richness and predicted richness showing similar results in the RF analysis (Table S5) and relatively lower in the *glm.nb* model. While these relationships are empirical, the strength of the predictions implies that, across the world's tropical forests, local species richness at the 1-ha scale is largely deterministic. The results showed that local tree diversity tends to converge on similar levels of richness when conditions are similar. This trend holds despite the fact that a minority of genera and almost no species of trees exist in common between the studied biogeographic regions, with different families and genera having undergone their own patterns of diversification within each region [29–32].

A noteworthy nuance emerges when comparing the relationship of individual environmental variables to species richness: although the correlations between any single environmental variable and richness are relatively weak, simultaneous incorporation of all environmental predictors in the RF analysis achieved high predictive power ($R^2 > 0.86$). Furthermore, sensitivity analysis highlighted that the importance of these interactions is context-dependent: the trained RF model reflected the specific range of environmental conditions in the current dataset. While predictive relationships might shift with the inclusion of additional forest plots (taking in a greater range of local environments or subregions), the sample size as it already stands is very large and in diverse settings, so we regard these inferences to be robust and unlikely to change substantially with the addition of more samples.

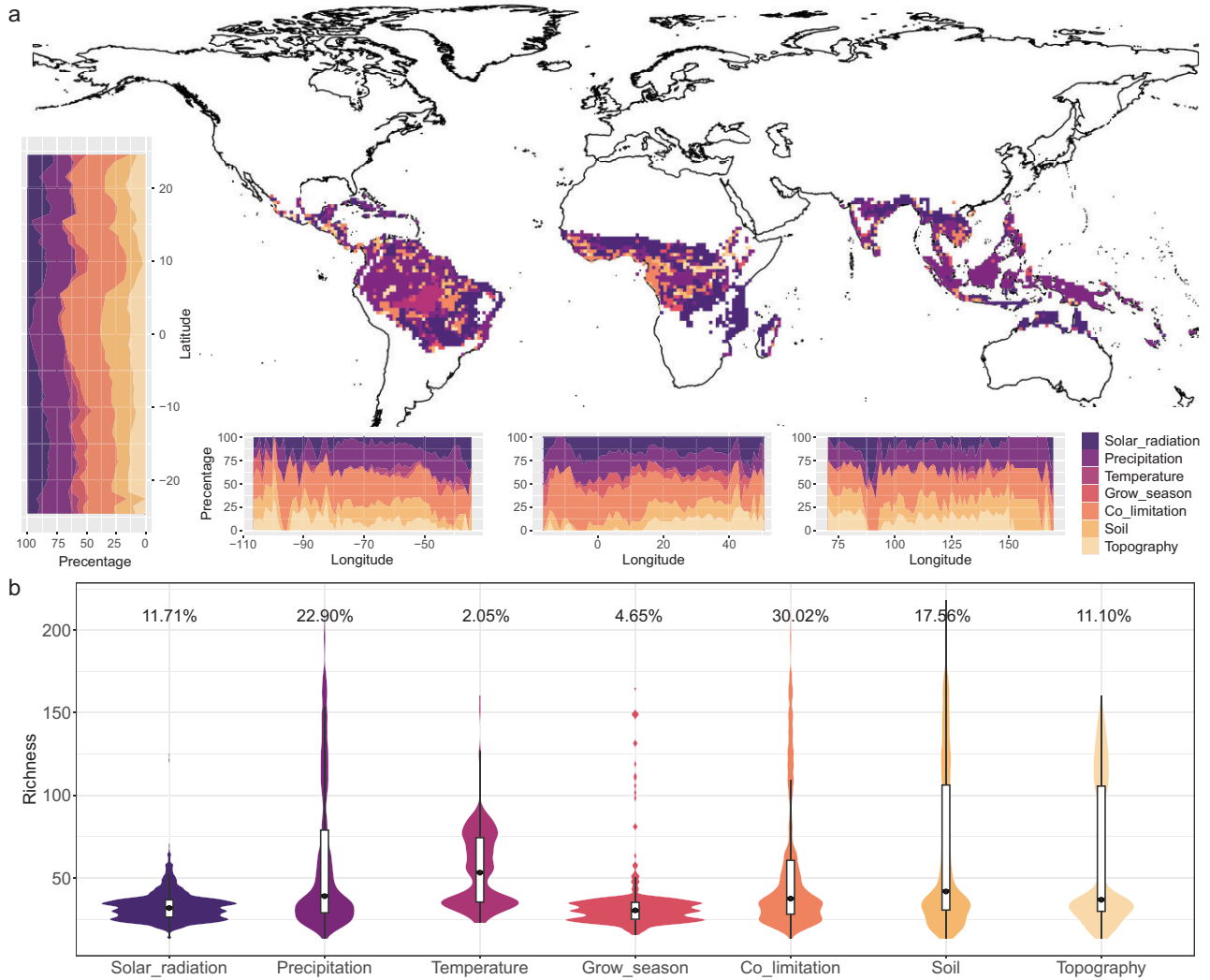


Figure 5. Geographical distribution of predominance of seven categories of environmental variables. Co-limitation refers to areas in which no single factor dominates. (a) Driving factors distribution pattern in pantropical regions, along longitude and latitude. (b) Percentage of main categories driving factors in tropical region. Detailed variables for different drivers are explained in Table S3. Review drawing number: GS京(2025) 2372号.

(Continued from previous page)

⁹⁰Department of Ecosystems and Conservation, Sokoine University of Agriculture, Morogoro Tanzania; ⁹¹Forest and Wildlife Research and Development Institute, Phnom Penh, Cambodia; ⁹²Department of Geography, School of Environment Education and Development, University of Manchester, Manchester M13 9PL, UK; ⁹³School of Biological Sciences, Universiti Sains Malaysia, Pulau Pinang 11800, Malaysia;

Our results suggest that moisture-related variables are more important at the large scale, while the soil-related variables play a relatively more important role at the local scale. The relatively strong local effect of soil factors is consistent with the mosaic of soil-influenced habitats that can exist at a fairly local scale within the tropics, producing differences in tropical tree communities [33,34].

The strongest climate-related predictors of tree richness identified in the RF and sensitivity analysis (Figs 4 and 5) were those reflecting aspects of the moisture supply to vegetation, with variations in temperature playing a less important role. The estimated npp itself is high on the list of predictors at large scales (50, 100 and 200 km; Fig. 4d

and Fig. S8). In general, the climate factors that emerge as importantly correlated with tree species richness were all consistent with conditions that favor plant growth and productivity rather than conditions that impose severe water stress or suppress photosynthesis due to low temperatures. This reinforces and refines the patterns linking tree diversity and plant physiology/productivity that have long been noted [1,2,6,35]. While relationships between species richness and climate have been demonstrated in other studies [36,37], what is striking in these new results is how predictable the overall relationship is when soil factors were included (discussed below) and how consistent the relationship is across different regions of the tropics.

(Continued from previous page)

⁹⁴Institute for Biodiversity and Environmental Research, Universiti Brunei Darussalam, Jalan Tungku Link, BE 1410, Brunei Darussalam; ⁹⁵Tropical Forest Research Institute, Jabalpur, Madhya Pradesh 482021, India; ⁹⁶Lab for Spatial Informatics, International Institute of Information Technology, Telangana 500032, India; ⁹⁷Centre for Environmental Studies, Amrita School of Engineering, Coimbatore, Amrita Vishwa Vidyapeetham 641112, India; ⁹⁸CIFOR-ICRAF, St Eugene Office Park, Lusaka, Zambia; ⁹⁹FORLIANCE GmbH, Bonn 53119, Germany; ¹⁰⁰State Key Laboratory of Plant Diversity and Specialty Crops, Xishuangbanna Tropical Botanical Garden, Chinese Academy of Sciences, Mengla 666303, China; ¹⁰¹Forest Research Institute, University of the Sunshine Coast, Sippy Downs, QLD, 4556, Australia; ¹⁰²Department of Science and Technology—Philippine Council for Industry, Energy and Emerging Technology Research and Development, Taguig City 1631, Philippines; ¹⁰³School of Environmental Sciences, Jawaharlal Nehru University, New Delhi 110067, India; ¹⁰⁴Department of Biology, Faculty of Sciences, National Polytechnic School of Ecuador, Quito 170525, Ecuador; ¹⁰⁵School of Biological and Environmental Sciences, Liverpool John Moores University, Liverpool L3 3AF, UK; ¹⁰⁶Department of Environment and Development Studies, United International University, Dhaka 1212, Bangladesh; ¹⁰⁷Tropical Forests and People Research Centre, University of the Sunshine Coast, Maroochydore DC, Queensland 4558, Australia; ¹⁰⁸Department of Climate Variability and Aquatic Ecosystems, Faculty of Ocean Science and Technology, Kerala University of Fisheries and Ocean Studies, Kerala 692508, India; ¹⁰⁹Environmental Studies and Research Center, Universidade Estadual de Campinas, UNICAMP, Campinas 13083-970, Brazil;

The question of why such environment–richness relationships exist for trees is one of the greatest conundrums in ecology and evolution [2]. Explanations vary from the legacy of a moist tropical origin of angiosperms [38] to the niche partitioning capabilities of more productive ecosystems [35] to the lack of physiological extremes, which allows functional equivalency and thus both greater diversity of growth architectures [39–41] and extensive niche overlap [3] that favors the co-existence of high species richness.

Certain internal patterns of influence on species richness within each tropical region are discernible (Fig. 5 and Figs S10–S13). Towards higher latitudes of the tropics, species richness was more strongly impacted by factors related to temperature and growing season, whereas, closer to the equator, moisture-related factors were more important (Fig. 5). Although particular categories of environmental factors cause a variation in the species richness in some areas, most regions are subject to co-limitation by multiple categories of environmental factors (Fig. 5). The high species richness observed in tropical regions may be maintained through the interaction of multiple factors and co-limitation [19]. In natural environments, plant growth is impacted by multiple variables [42], with the optimal growth occurring when these interacting variables attain a state of conducive equilibrium [43]. In this study, these co-limitation areas likely create environmental conditions under which resources are relatively balanced, favoring the survival of various species and thus making these areas high in species richness. Given the relatively limited attention given to the role of soil parameters in affecting the patterns of tropical tree diversity, soil pH surprisingly emerges as influential in the RF analysis, especially at the local scale (Fig. 4b and c, and Fig. S8). The role of soil pH has long been known as an influence on more local-scale patterns in tropical forest composition [34] and soil pH is known to play a major role in plant ecology in general [44]. Soil pH correlates with a range of other factors, such as nutrient concentrations and availability, including the mobility of toxic ions such as aluminum [45]. The species composition of vegetation can itself influence the soil pH and other soil factors [46–48], but the possibility of complex feedback loops between species diversity and soil being significant on the pantropical scale can only be considered as speculative.

While a lower soil pH may be seen as a physiologically more extreme environment, for the above reasons, in fact, tree species richness tends

to be highest in the lowest pH soils in our dataset (Figs S9b and S10b). Empirically, this supports previous results from studies in other biomes that plant species richness peaks in low soil pH of ~4 [49] and declines as the pH increases (Fig. S9b). A widely discussed principle in plant ecology is that a certain degree of physiological ‘stress’—such as low soil pH—may suppress plant growth and productivity, reducing the competitive ability of faster-growing generalists [44]. In the context of disturbance events at varying scales, this may reduce competitive exclusion and enable a greater number of species to coexist [44,50,51].

While it is reasonable to focus on the potential effect of environmental factors on the trees themselves, it is also crucial to consider that other factors could be at work, without directly involving the physiology of the trees as the primary driver. For instance, the Janzen–Connell hypothesis suggests that the diversity levels in tropical forests are controlled by the intensity of attacks by insect herbivores and pathogens [2,52–54]—with constantly warm and moist conditions favoring the specialization of insect pest or pathogen populations and the strong density-dependent control of tree populations allowing more tree species to coexist locally. There is considerable evidence that a degree of selective pest pressure can maintain diversity in plant communities, but inconsistent evidence that density-dependent mortality is stronger in the tropics than in temperate regions [2]. However, to explain the patterns seen here, the pest-pressure effect would need to operate in a finely modulated way along environmental and species richness gradients within the tropics, quite aside from whether it differs between tropical and temperate regions. A potential effect of the soil pH on herbivory—perhaps with respect to the extent of the species pools on different soil types and perhaps in affecting the nutrient or secondary compound content of plants or the growth rate of fresh edible tissues—might also be involved in the observed relationship with soil factors.

Convergence in richness amongst biogeographic regions

From the analysis of this dataset, there is no obvious evidence for any regional influence producing anomalously high or low tree species richness relative to the overall pantropical trend. When 1-ha richness data from all the regions are overlaid on the same scatterplot, the Americas, Asia (including Wallacea and Australasia) and tropical Africa all fall close to one another, within

margins of error (Fig. 3 and Figs S14–S16). The spatial RF model demonstrates that analogous environments support comparable richness levels, irrespective of regional species pools (Fig. S17). This contrasts with the overall comparison of regions without accounting for equivalent environments. Most strikingly, there is particularly high species richness for some plots in the Americas (Fig. 3 and Figs S14–S16). One of the primary predictors, and possible drivers, of the elevated diversity in the neotropics appears to be this region's consistently higher levels of moisture, as expressed by a range of factors in the RF model [55]. This may be further reinforced by the region's relatively low soil pH values, which we find correlate significantly with species richness independently of climate. When the RF model was implemented, the anomalously high richness of the Americas disappeared and was explicable almost exclusively in terms of empirical environmental factors.

The close similarity in richness between tropical regions, when compared in terms of the combinations of environmental factors that best predict richness, is evident despite the divergent geological, climatic and evolutionary histories of these regions that span tens of millions of years [29,38,56,57]. Independent evolution and diversification of clades has given a distinctive taxonomic composition to the forests in each of the main regions—the Americas, South East Asia (and Wallacea/Oceania) and tropical Africa [38,56,57]. In each region, the composition and distribution of flora have been influenced by continental collisions, mountain building and the cooling and drying of the climate in the Cenozoic [30,56–59]. For instance, tropical Africa lost a large proportion of its previous diversity during the late Cenozoic/Quaternary due to the drying of the climate, especially during glacial episodes [38,56,57], and now has restricted total richness and phylogenetic diversity of trees at the regional scale [23,30]. The fossil pollen record also shows that the rainforests of northern Australia have been through strong drying episodes in which their areal extent was severely restricted [38,56,57]. While the climate history of the tropical forests of India is not well known, it is possible that, as a relatively dry and small rainforest enclaves it, it could also have gone through past climate bottlenecks [56]. The uplift of the Andes created a complex topographical gradient that facilitated allopatric speciation and the establishment of numerous microhabitats, leading to species diversification [60,61].

Despite all of these different histories and potential trajectories, brought about by climatic and tectonic history, and regional-scale diversifica-

tions, it is striking that, from the perspective of the 1-ha-sample scale, all the regions that we distinguish here adhere closely to the same pantropical pattern of richness in relation to present-day environmental conditions (Fig. 2 and Figs S14–S16).

The very close correspondence in tree species richness between the regions, despite all of the historical legacy factors that could have potentially caused a divergence in diversity, implies the existence of precise control by 'governance' factors in the forest community that tends to cause the richness to settle at a particular level. Many potential mechanisms have been put forward to explain how the tree species richness in tropical forests is maintained [2], including those discussed above. It is, however, surprising that the mechanisms at work are able to operate so precisely, all across the tropics, to modulate the local-scale richness when there are so many factors that would be expected to cause the richness to diverge. Overall, the mechanisms invoked to explain the high tree species richness of tropical forests and its variation within the world's tropical forests can be grouped into two kinds. Disequilibrium hypotheses invoke time-dependent processes of the progressive buildup of diversity by diversification or migration and its destruction by extinction [58]. According to such mechanisms, there is no 'lid' on the maximum diversity in the tropics and differences in diversity reflect the balance between diversification events and extinction events. Equilibrium hypotheses, by contrast, assume that there is a set capacity to the number of tree species that can co-exist in any one place and that differences in diversity reflect differences in this capacity. This set of mechanisms necessarily depends upon differences in niche structuring—e.g. the number of discrete niches available due to the heterogeneity of microenvironments [62], the narrowness of specialized niches that is possible in a given environment (affecting the opportunities for slotting in extra species) [63] or the degree of overlap in tree species niches that can occur before competitive exclusion begins to reduce the diversity [64,65]. In our opinion, the results of this study support a predominance of equilibrium- or niche-based mechanisms, as the relationships between environmental parameters and hectare-scale species richness are so strongly convergent between different parts of the world. If the vagaries of diversification and extinction were more important in affecting the richness, then we might expect to see large differences in local diversity between different regions under similar environmental conditions.

Whatever the ecological mechanisms that mediate the relationship between environmental

(Continued from previous page)

¹¹⁰Research Centre in Botany, VHNSN College, Virudhunagar 626001, India; ¹¹¹Department of Forestry, Uttar Banga Krishi Viswavidyalaya, Pundibari, Cooch Behar, West Bengal 736165, India; ¹¹²Department of Botany, Institute of Biology, Federal University of Rio de Janeiro, Rio de Janeiro 21941-590, Brazil; ¹¹³IBAM—Instituto Bem Ambiental, Belo Horizonte 30130-090, Brazil; ¹¹⁴Myr Projetos Sustentáveis (Grupo Myr), Belo Horizonte 30130-009, Brazil; ¹¹⁵Bell Museum, University of Minnesota, St Paul, MN 55108, USA; ¹¹⁶State Key Laboratory of Oil and Gas Reservoir Geology and Exploitation & Institute of Sedimentary Geology, Chengdu University of Technology, Chengdu 610059, China; ¹¹⁷Naturalis Biodiversity Center, 2300 RA, Leiden, The Netherlands; ¹¹⁸CAVElab -Computational & Applied Vegetation Ecology, Department of Environment, Faculty of Bioscience Engineering, Ghent University, 9000 Gent, Belgium; ¹¹⁹Escuela Nacional de Estudios Superiores, Universidad Nacional Autónoma de México, Mérida, 97357, Mexico; ¹²⁰Conservation Ecology Program, King Mongkut's University of Technology Thonburi, Bangkok 10150, Thailand; ¹²¹Centre for Tropical Environmental and Sustainability Science, and College of Science and Engineering, James Cook University, Cairns, Queensland 4878, Australia; ¹²²Programa de Maestría en Manejo de Bosques, Universidad de Los Andes, Via Chorro de Milla, Mérida 5101, Venezuela and ¹²³Southeast Asia Biodiversity Research Institute, Chinese Academy of Sciences & Xishuangbanna Tropical Botanical Garden, Chinese Academy of Sciences, Mengla 666303, China.

parameters and tree species richness, they appear to operate consistently and in combination along a sliding scale in terms of the levels of richness that they permit. Whilst it is intuitively hard to accept that such mechanisms could exert their effects so precisely between different regions of the world with their own distinct tropical tree floras that have been separated for tens of millions of years, this is indeed what the results of our study suggest.

CONCLUSIONS

Although regional species pools differ due to distinct geological histories, analogous environmental conditions yield similar local richness patterns. The broad-scale approach employed in this study, combined with spatial RF analysis, has demonstrated that there is a striking predictability in pantropical tree species richness sampled at the hectare scale. This predictability involves a combination of environmental factors, with climate and to some extent soil showing strong correlations with tree species richness patterns across and between tropical regions. Pantropical regions exhibit high species richness, primarily due to the intricate interplay of co-limitation that creates stable and balanced conditions. Although regional species pools differ, analogous environmental conditions yield similar local richness patterns. A multifactorial interplay of evolutionary and ecological mechanisms is presumably at work in controlling this consistency of species richness, preventing regional divergence. The observed variation in tropical species richness must be seen as the outcome of a range of different factors acting simultaneously, sometimes in parallel and sometimes in opposition to one another. It is possible that these multiple factors each act through a range of different ecological mechanisms, requiring separate elucidation. These findings decisively strengthen the case for the niche theory, revealing its remarkable predictive power. They provide compelling evidence that niche segregation is a pervasive and dominant force, structuring communities from local patches to broad regional landscapes, even in the presence of diverse regional species assemblages. Furthermore, by elucidating the contemporary environmental controls on species richness, this work contributes a vital context for predicting plant diversity changes under future warming scenarios.

Whilst identification of the true underlying mechanisms involved remains a fundamental challenge for ecology, this study contributes to the ongoing challenge in ecology to iden-

tify and understand the controls on biological diversity.

In this paper, we had originally hypothesized that local-scale tree α diversity across the global tropics can be consistently predicted by using contemporary environmental variables, such that analogous environmental conditions will support similar levels of richness, irrespective of deep historical biogeographic divergences. This hypothesis has survived its test, with striking predictability of species richness as sampled at the hectare scale. We also hypothesized that variation in richness on more localized scales between samples would be governed by a distinct set of influences. Despite some subtle scale-related differences, this was essentially disproven, with similar sets of environmental factors governing throughout.

It is necessary to keep in mind, however, that our study is confined to data obtained at the 1-ha sampling scale and that other patterns may emerge at other local sampling scales [10] or for other life forms (e.g. lianas)—either in sub-hectare or larger plots, or at the level of beta-diversity turnover rates. This awaits other studies and the additional information that comes from these will shed more light on the mechanisms at work behind tropical tree diversity patterns. The findings of this study are derived from current sample surveys and the uneven distribution of these samples may have affected our results. We advocate for an increase in sample surveys in tropical regions and the development of a more comprehensive tropical sample database in future scientific research.

Intriguingly, other aspects of the community structuring of tropical forests may be found to show striking convergence patterns across the tropics. Cooper *et al.* [32] have recently revealed a strong convergence in relative abundance data—rather than richness, as is shown here—in forest plots throughout the tropics. There is a need for further careful comparisons of the structure and functioning of tropical forests across different regions in order to understand how closely they have maintained their similarities and to better understand the driving mechanisms behind the observed patterns.

MATERIALS AND METHODS

Tree data and richness

Tree inventory data were assembled from the PFN of old-growth (not recently logged or cleared) closed canopy forest plots from across the global tropics, including the tropical dry forest and its various transitional forms to the tropical

rainforest [23]. All trees, defined as free-standing woody individuals (including palms) with a diameter at breast height (1.3 m) of ≥ 10 cm, were measured and identified in each plot (Data S1). If the species names could not be determined, then plot-specific morphospecies were recorded, with the closest taxonomic assignment. All morphotypes included here were identified to the Linnean species, genus or at least family level. Unknowns at the Linnean family level were not included.

The numbers of individuals and species were obtained from a total of 429 1-ha plots located in the Americas (197 plots), Africa (150 plots) and Asia (82 plots) (Fig. 2, Figs S3 and S4, Data S2). The Asian region included Wallacea and Australasia based on a shared floristic affinity [23]. The geographical boundaries of the tropics were defined based on the map in [66].

In addition to the observed species richness, we analysed the richness in each 1-ha plot estimated by using sample-based rarefaction and Fisher's α (see Methods S1). Both of these methods are widely recognized and frequently used measures of species richness [24]. Sample-based rarefaction is often considered a more reliable estimator of true species richness in a community, particularly when sampling is incomplete, as is frequently the case in diverse assemblages [25,67]. This technique allows a more accurate comparison across communities by taking into account undetected species and different levels of sampling effort. We used the function 'estimated' from the iNEXT package in R [68] to estimate species richness based on sample coverage, which accounts for the completeness of sampling by estimating how well a community has been sampled. The richness was standardized to the same sample coverage (1 ha) to compare the survey completeness of each sample size, providing a robust measure that accounts for uneven sampling efforts across the regions. The rarefaction curves were plotted to visualize the richness of species across the coverage sampling efforts, indicating that the sample coverage values for all the localities are quite high—most of them are >0.8 (Fig. S1).

Our primary analysis focused on species richness estimated from sample-based rarefaction to ensure the robustness of our estimates. Fisher's α was also examined given its applicability to communities in which the species follow a log-series pattern with high proportions of rare species and for detecting the influence of abundance distributions on species diversity [26,69]. Compared with other diversity measures, these two indices showed the strongest correlation with the ob-

served richness (Table S4 and Fig. S2; see also Methods S1).

Environmental data

We collected data surfaces for an initial set of 65 environmental predictors, including bioclimatic, soil and topographic variables (Data S3–S5). Bioclimatic data were sourced from CHELSA (Climatologies at high resolution for the Earth's Land Surface Areas; <http://chelsa-climate.org/>), which provides climate data at a spatial resolution of 30 arc-seconds (~ 1 km²). Soil data were obtained from the ISRIC World Soil Information SoilGrids dataset (<https://data.isric.org/>), which provides model-interpolated predictions of soil parameters at a resolution of 250 m. These predictions are derived from the integration of data from thousands of soil cores from across the globe, with geological, surface sediment, topographic, normalized difference vegetation index and climatic background information [15]. Topographic data with a spatial resolution of ~ 1 km were downloaded from EarthEnv (<http://www.earthenv.org/topography>). Important environmental variables were selected by using the Boruta algorithm from the 'Boruta' package [70] in R (Data S6). Multicollinearity was addressed by assessing the correlation and variance inflation factors, resulting in a refined set of 24 environmental predictors for the RF analysis (Table S1).

Spatial RF

To predict species richness based on multiple environmental variables, we applied a RF combined with spatial regression by using the 'spatialRF' package in R [22], which can be applied on regular or irregular data [71]. This package enhances traditional RF techniques by accounting for spatial autocorrelation when observations are not independent, but rather show geographic relationships. Spatial autocorrelation based on Moran's I index was explicitly taken into account to improve the model accuracy in capturing spatial patterns. To ensure robust evaluation of the model, we employed cross-validation by dividing the data into 30 spatial folds for training and testing. Additionally, spatially thinned occurrence data were generated by using the 'thin' function from the 'spThin' package [72], which filters occurrence locations to ensure they are a set minimum distance apart (e.g. 50 km). This spatial thinning reduces bias from uneven species collections and was used in RF analysis to compare the results from the full dataset (see Methods S2).

Statistical analysis

We used linear regression to examine the relationship between the species richness predicted by using non-spatial and spatial RF and that estimated from the sample-based rarefaction, Fisher's α and observed richness. The PCA was performed to explore variation between the 24 environmental variables in the three major tropical regions: the Americas, Africa and Asia. We used the 'FactoMineR' package [73] in R to implement the PCA and the 'factoextra' package [74] to generate a biplot that enabled visualization of the results.

Spatial RF excels at capturing complex, nonlinear relationships and identifying important predictors, primarily due to its predictive power [75]. As a complementary approach, the Generalized Linear Model with Negative Binomial distribution (glm.nb) was used. This allowed us to validate the RF results by statistically testing broad-scale factors such as continental differences and examining linear relationships, combining data-driven prediction with statistical inference. To quantify the effects of the 24 environmental variables on the species richness across the three tropical regions, we applied a negative binomial generalized linear model (glm.nb) by using the 'MASS' package in R and compared the results with those from spatial RF. Spatial RF excels in capturing complex, nonlinear relationships between species richness and environmental variables, but does not provide interpretable coefficients that explain the direction and magnitude of those relationships [76]. As a complementary approach, glm.nb was used to validate the robustness of the RF results by enabling the identification and interpretation of regional differences and linear relationships between environmental predictors and species richness. Both the main effects and the interactions between environmental variables and regions were examined by using this approach (see [Methods S3](#)).

Sensitivity analysis

To determine which class of environmental variables (e.g. precipitation, temperature, soil, etc.) best explained the variation in the local tree species richness, we classified the 24 environmental variables into seven categories—temperature, precipitation, growing season, solar radiation, soil, topography and co-limitation—and conducted a sensitivity analysis of the RF model by following the methods of Saltelli *et al.* [77] and Liang *et al.* [19]. All of the above analyses were performed in R version 4.2.3 (R Core Team, 2024).

The R script used for all analyses is provided in [Data S6](#).

We conducted a sensitivity analysis through the following steps:

Step 1: Using all environmental variables $X(s)$, we applied the RF model to simulate the predicted species richness $Y_{all}(s)$:

$$Y_{all}(s) = f(X(s)),$$

where $f()$ represents the RF model, $X(s)$ represents the values of the environmental variables and s represents the six categories to which the environmental variables belong: E1, temperature; E2, precipitation; E3, growing season; E4, solar radiation; E5, soil; E6, topography ([Table S3](#)).

Step 2: We then applied the RF model to predict the tree species richness based on all the environmental variables except those belonging to E1, $S_{-E1}(s)$:

$$Y_{-E1}(s) = f_{-E1}((X - E1)(s)),$$

where $f_{-E1}()$ represents the RF model simulated with all the variables except those associated with temperature and $(X - E1)(s)$ represents the variables of the remaining five categories (E2–E6).

Step 3: We calculated the relative sensitivity of the predicted species richness to E1 from:

$$R(E1) = |Y_{all}(s) - Y_{-E1}(s)| / Y_{all}(s).$$

Step 4: We repeated Steps 2 and 3 to calculate the relative sensitivity of each of the remaining categories E2–E6. For a given area, the category with the highest relative sensitivity and meeting the threshold of relative sensitivity of $\geq 1/7$ was considered that which best explained the variation in the tree richness for that area.

Step 5: In areas in which the relative sensitivities were $< 1/7$ for all the categories, we hypothesized that the tree richness was not related to any single category, but rather multiple categories of environmental variables. Therefore, we created a seventh category (E7) called co-limitation to characterize areas in which no single type of environmental factor dominates.

Step 6: Steps 1–5 were repeated to calculate the relative sensitivity of each of the seven categories, including E7. To visualize the regional variation in category importance, we calculated the relative sensitivity of each category as a percentage of that of all the categories and plotted

this along latitudinal and longitudinal gradients spanning the tropics on a map.

DATA AVAILABILITY

All datasets, R scripts and model output results have been uploaded to Figshare ([Data S1–S7](#) and Appendix in [Supplementary materials](#)).

SUPPLEMENTARY DATA

Supplementary data are available at [NSR](#) online.

ACKNOWLEDGEMENTS

We thank Akihiro Nakamura and Yishan Zhao from Xishuangbanna Tropical Botanical Garden, Chinese Academy of Sciences, for their assistance and insightful suggestions on data analysis and figure color selection during the revision process, and Dr Binu Tripathi (West Virginia University) for help with data formatting. We thank the reviewers for their very thorough and helpful comments.

FUNDING

This work was supported by the National Key R&D Program of China (2022YFF0800800), the National Natural Science Foundation of China (NSFC) (42320104005, 42372033 and 32372058), the Young and Middle-Aged Academic and Technical Leaders of Yunnan (202305AC160051), the 14th Five-Year Plan of the Xishuangbanna Tropical Botanical Garden, Chinese Academy of Sciences (E3ZKFF7B) and the Southeast Asia Biodiversity Research Institute, Chinese Academy of Sciences (Y4ZK111B01).

AUTHOR CONTRIBUTIONS

S.M.X. conducted analysis, writing and interpretation. J.M.A. initiated and conceptualized the study, conducted interpretation and led the writing. S.F.L. conducted analysis, writing and interpretation. F.S. coordinated data organization and assisted in interpretation. K.W.T. assisted in data interpretation. D.M.G., T.S. and Z.K.Z. assisted with checking and editing the manuscript. PFN members A.Q., A.S., A.H.R., A.A.M., A.R.M., A.G., A.A.D., A.M., A.N., B.H., C.A. J., D.P., D.J.H., D.R.D., D.S., D.S.B., E.L.de.O., E.N., E.van.den.B., E.L.W., F.H.I., F.Z.S., F.W., F.M.A., G.D., G.S., G.I-M., H.N., I.T., I.C.Z-B., I.H-C., J.O.R-C., J.G., J.A.D., J.S., J.H., J.R.P., J.N.W., J.L., J.D.D.A., J.R.L., J.R.R.P., J.E.G-A., J.C-A., J.H., K.B-G., K.R.H., K.F., K.W.T., L.R., L.O.D., L.Y., L.F.A., L.T.M., M.S., M.J.M., M.T., M.S.S., M.A.R.P., M.T.F.P., M.S., M.van.de.B., M.L.B., M.L.K., M.S.H., M.K., M.J.L., M.A.M-R., N.P., N.S., N.B.K., N.J., N.P., O.O., O.R.V., P.M., P.C., P.da.C.B., R.Z., R.S.S., R.M., R.C.P., R.V.S., R.D.H., R.K.V., R.S., R.L.C., R.T., S.G.R., S.C.G., S.B., S.W., S.A.M., S.J., S.A.V., S.M., S.S., S.C., S.K., T.W., T.M., T.W., V.A.T.R., V.H., V. A-R., W.C., W.F.L. and Y.A.C. contributed research data. Many coauthors made helpful suggestions for the analysis and writing.

Conflict of interest statement. None declared.

REFERENCES

- Gentry AH. Changes in plant community diversity and floristic composition on environmental and geographical gradients. *Ann Missouri Botanical Garden* 1988; **75**: 1–34.
- Hill JL and Hill RA. Why are tropical rain forests so species rich? Classifying, reviewing and evaluating theories. *Prog Phys Geogr* 2001; **25**: 326–54.
- Hubbell SP. *The unified neutral theory of biodiversity and biogeography*. New Jersey: Princeton University Press, 2001.
- Adams JS. *Species Richness: Patterns in the Diversity of Life*. Berlin, New York and Chichester, UK: Springer, in association with Praxis, 2009.
- Houlahan JE, McKinney ST, Anderson TM *et al*. The priority of prediction in ecological understanding. *Oikos* 2017; **126**: 1–7.
- Hawkins BA, Albuquerque FS, Araújo MB *et al*. A global evaluation of metabolic theory as an explanation for terrestrial species richness gradients. *Ecology* 2007; **88**: 1877–88.
- Ricklefs RE and He F. Region effects influence local tree species diversity. *Proc Natl Acad Sci USA* 2016; **113**: 674–9.
- Ter Steege H, Pitman NCA, Do Amaral IL *et al*. Mapping density, diversity and species-richness of the Amazon tree flora. *Commun Biol* 2023; **6**: 1130.
- Parmentier I, Malhi Y, Senterre B *et al*. The odd man out? Might climate explain the lower tree α -diversity of African rain forests relative to Amazonian rain forests? *J Ecol* 2007; **95**: 1058–71.
- Keil P and Chase JM. Global patterns and drivers of tree diversity integrated across a continuum of spatial grains. *Nat Ecol Evol* 2019; **3**: 390–9.
- Chu C, Lutz JA, Král K *et al*. Direct and indirect effects of climate on richness drive the latitudinal diversity gradient in forest trees. *Ecol Lett* 2019; **22**: 245–55.
- Chase JM, McGill BJ, Thompson PL *et al*. Species richness change across spatial scales. *Oikos* 2019; **128**: 1079–91.
- ForestPlots.net, Blundo C, Carilla J, Grau R *et al*. Taking the pulse of Earth's tropical forests using networks of highly distributed plots. *Biol Conserv* 2021; **260**: 108849.
- Karger DN, Conrad O, Böhner J *et al*. Climatologies at high resolution for the earth's land surface areas. *Sci Data* 2017; **4**: 170122.
- Hengl T, Jesus JMd, Heuvelink GBM *et al*. SoilGrids250m: global gridded soil information based on machine learning. *PLoS One* 2017; **12**: e0169748.
- Amatulli G, Domisch S, Tuanmu M-N *et al*. A suite of global, cross-scale topographic variables for environmental and biodiversity modeling. *Sci Data* 2018; **5**: 180040.
- Poggio L, de Sousa LM, Batjes NH *et al*. SoilGrids 2.0: producing soil information for the globe with quantified spatial uncertainty. *SOIL* 2021; **7**: 217–40.
- Brun P, Zimmermann NE, Hari C *et al*. Global climate-related predictors at kilometer resolution for the past and future. *Earth Syst Sci Data* 2022; **14**: 5573–603.

19. Liang J, Gamarra JG, Picard N *et al.* Co-limitation towards lower latitudes shapes global forest diversity gradients. *Nat Ecol Evol* 2022; **6**: 1423–37.
20. Cai L, Kreft H, Taylor A *et al.* Global models and predictions of plant diversity based on advanced machine learning techniques. *New Phytol* 2023; **237**: 1432–45.
21. Wright MN and Ziegler A. Ranger: a fast implementation of random forests for high dimensional data in C++ and R. *J Stat Soft* 2017; **77**: 1–17.
22. Benito BM. spatialRF: easy spatial modeling with random forest. 2022.
23. Slik JWF, Arroyo-Rodriguez V, Aiba S-I *et al.* An estimate of the number of tropical tree species. *Proc Natl Acad Sci USA* 2015; **112**: 7472–7.
24. Gotelli NJ and Colwell RK. Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. *Ecol Lett* 2001; **4**: 379–91.
25. Chao A and Jost L. Coverage-based rarefaction and extrapolation: standardizing samples by completeness rather than size. *Ecology* 2012; **93**: 2533–47.
26. Fisher RA, Corbet AS, Williams CB. The relation between the number of species and the number of individuals in a random sample of an animal population. *J Anim Ecol* 1943; **12**: 42–58.
27. Capinha C, Essl F, Seebens H *et al.* Models of alien species richness show moderate predictive accuracy and poor transferability. *NB* 2018; **38**: 77–96.
28. Shafullah AZ, Werner J, Kennedy E *et al.* Machine learning based prediction of insufficient herbage allowance with automated feeding behaviour and activity data. *Sensors* 2019; **19**: 4479.
29. Slik JWF, Franklin J, Arroyo-Rodríguez V *et al.* Phylogenetic classification of the world's tropical forests. *Proc Natl Acad Sci USA* 2018; **115**: 1837–42.
30. Qian H, Kessler M, Zhang J *et al.* Angiosperm phylogenetic diversity is lower in Africa than South America. *Sci Adv* 2023; **9**: ead1022.
31. Mukul SA, Herbohn J, Firn J. Rapid recovery of tropical forest diversity and structure after shifting cultivation in the Philippines uplands. *Ecology and Evolution* 2020; **10**: 7189–211.
32. Cooper DLM, Lewis SL, Sullivan MJP *et al.* Consistent patterns of common species across tropical tree communities. *Nature* 2024; **625**: 728–34.
33. Whitmore TC. *An introduction to tropical rain forests*. Oxford: Oxford University Press. 1990.
34. Pires J and Prance G. The vegetation types of the Brazilian Amazon. 1985.
35. Currie DJ and Paquin V. Large-scale biogeographical patterns of species richness of trees. *Nature* 1987; **329**: 326–7.
36. Saupe EE, Myers CE, Townsend Peterson A *et al.* Spatio-temporal climate change contributes to latitudinal diversity gradients. *Nat Ecol Evol* 2019; **3**: 1419–29.
37. Coelho MTP, Barreto E, Rangel TF *et al.* The geography of climate and the global patterns of species diversity. *Nature* 2023; **622**: 537–44.
38. Morley RJ. *Origin and evolution of tropical rain forests*. West Sussex: John Wiley & Sons, 2000.
39. Steenis C. *Patio ludens and extinction of plants. Notes from the Royal Botanic Garden Edinburgh* 1978; 317–23.
40. Kleidon A, Adams J, Pavlick R *et al.* Simulated geographic variations of plant species richness, evenness and abundance using climatic constraints on plant functional diversity. *Environ Res Lett* 2009; **4**: 014007.
41. Reu B, Proulx R, Bohn K *et al.* The role of climate and plant functional trade-offs in shaping global biome and biodiversity patterns. *Global Ecol Biogeogr* 2011; **20**: 570–81.
42. Domingues TF, Meir P, Feldpausch TR *et al.* Co-limitation of photosynthetic capacity by nitrogen and phosphorus in West Africa woodlands. *Plant Cell & Environment* 2010; **33**: 959–80.
43. Chapin FS, Bloom AJ, Field CB *et al.* Plant responses to multiple environmental factors. *Bioscience* 1987; **37**: 49–57.
44. Grime JP. *Vegetation Processes, and Ecosystem Properties*. West Sussex: John Wiley & Sons, 2006.
45. van Breemen N and Finzi AC. Plant-soil interactions: ecological aspects and evolutionary implications. *Biogeochemistry* 1998; **42**: 1–19.
46. Kulmatiski A, Beard KH, Stevens JR *et al.* Plant–soil feedbacks: a meta-analytical review. *Ecol Lett* 2008; **11**: 980–92.
47. van der Putten WH, Bardgett RD, Bever JD *et al.* Plant–soil feedbacks: the past, the present and future challenges. *J Ecol* 2013; **101**: 265–76.
48. Zhao S and Riaz M. Plant–soil interactions and nutrient cycling dynamics in tropical rainforests. In: Fahad S, Saud S, Nawaz T, Gu L, Ahmad M, Zhou R (eds.). *Environment, Climate, Plant and Vegetation Growth*. Cham: Springer Nature Switzerland, 2024, 229–64.
49. Azevedo LB, Van Zelm R, Hendriks AJ *et al.* Global assessment of the effects of terrestrial acidification on plant species richness. *Environ Pollut* 2013; **174**: 10–5.
50. Huston MA. *Biological Diversity: the Coexistence of Species*. Cambridge: Cambridge University Press, 1994.
51. Phillips OL, Hall P, Gentry AH *et al.* Dynamics and species richness of tropical rain forests. *Proc Natl Acad Sci USA* 1994; **91**: 2805–9.
52. Janzen DH. Herbivores and the number of tree species in tropical forests. *Am Nat* 1970; **104**: 501–28.
53. Connell JH. On the role of natural enemies in preventing competitive exclusion in some marine animals and in rain forest trees. In: Den Boer, P.J. and Gradwell, G.R. (eds.). *Dynamics of Populations*. Wageningen: Centre for Agricultural Publishing and Documentation. 1971; **298**: 1–610.
54. Bagchi R, Gallery RE, Gripenberg S *et al.* Pathogens and insect herbivores drive rainforest plant diversity and composition. *Nature* 2014; **506**: 85–8.
55. Raven PH, Gereau RE, Phillipson PB *et al.* The distribution of biodiversity richness in the tropics. *Sci Adv* 2020; **6**: eabc6228.
56. Morley RJ and Morley HP. The prelude to the Holocene: tropical Asia during the Pleistocene. In: N. Mehrotra, R. Asthana and S. Gupta (eds.). *Holocene Climate Change and Environment* Amsterdam: Elsevier, 2022: 1–32.
57. Morley RJ. The establishment of Palaeotropical rainforests from Africa to Oceania in relation to plate tectonics and zonal tropical climates. *SP* 2025; **549**: 113–36.
58. Hagen O, Skeels A, Onstein RE *et al.* Earth history events shaped the evolution of uneven biodiversity across tropical moist forests. *Proc Natl Acad Sci USA* 2021; **118**: e2026347118.
59. Currano ED, Jacobs BF, Pan AD. Is Africa really an 'odd man out'? Evidence for diversity decline across the Oligocene-Miocene boundary. *Int J Plant Sci* 2021; **182**: 551–63.
60. Antonelli A, Nylander JAA, Persson C *et al.* Tracing the impact of the Andean uplift on Neotropical plant evolution. *Proc Natl Acad Sci USA* 2009; **106**: 9749–54.
61. Hoon C, Wesselingh FP, ter Steege H *et al.* Amazonia through time: Andean uplift, climate change, landscape evolution, and biodiversity. *Science* 2010; **330**: 927–31.
62. Svenning J-C. On the role of microenvironmental heterogeneity in the ecology and diversification of Neotropical rain-forest palms (Arecaceae). *Bot Rev* 2001; **67**: 1–53.

63. Polato NR, Gill BA, Shah AA *et al.* Narrow thermal tolerance and low dispersal drive higher speciation in tropical mountains. *Proc Natl Acad Sci USA* 2018; **115**: 12471–6.
64. Levi T, Barfield M, Barrantes S *et al.* Tropical forests can maintain hyperdiversity because of enemies. *Proc Natl Acad Sci USA* 2019; **116**: 581–6.
65. Freeman BG, Strimas-Mackey M, Miller ET. Interspecific competition limits bird species' ranges in tropical mountains. *Science* 2022; **377**: 416–20.
66. Rubel F and Kottek M. Observed and projected climate shifts 1901–2100 depicted by world maps of the Köppen-Geiger climate classification. *Meteorologische Zeitschrift* 2010; **19**: 135–41.
67. Jost L. The relation between evenness and diversity. *Diversity* 2010; **2**: 207–32.
68. Hsieh T, Ma K, Chao A. iNEXT: an R package for rarefaction and extrapolation of species diversity (Hill numbers). *Methods Ecol Evol* 2016; **7**: 1451–6.
69. McGill BJ, Etienne RS, Gray JS *et al.* Species abundance distributions: moving beyond single prediction theories to integration within an ecological framework. *Ecol Lett* 2007; **10**: 995–1015.
70. Kursa MB and Rudnicki WR. Boruta: Wrapper Algorithm for All Relevant Feature Selection 2009:8.0.0. <https://doi.org/10.32614/CRAN.package.Boruta>.
71. Breiman L. Random forests. *Machine Learning* 2001; **45**: 5–32.
72. Aiello-Lammens ME, Boria RA, Radosavljevic A *et al.* spThin: Functions for Spatial Thinning of Species Occurrence Records for Use in Ecological Models 2014:0.2.0. <https://doi.org/10.32614/CRAN.package.spThin>.
73. Lê S, Josse J, Husson F. FactoMineR: an R package for multivariate analysis. *J Stat Soft* 2008; **25**: 1–18.
74. Kassambara A and Mundt F. *Factoextra: Extract and Visualize the Results of Multivariate Data Analyses*. CRAN, Contributed Packages, 2016.
75. Benito B. SpatialRF: easy spatial regression with random forest. *R Package Version* 2021; **1**.
76. Elmotawakkil A, Sadiki A, Enneya N. Predicting groundwater level based on remote sensing and machine learning: a case study in the Rabat-Kénitra region. *J Hydroinf* 2024; **26**: 2639–67.
77. Saltelli A, Ratto M, Andres T *et al.* *Global Sensitivity Analysis: the Primer*. West Sussex: John Wiley & Sons, 2008.