Wich, S

Analogous Environments Across the Tropics Have 2 Similar Levels of Tree Species Alpha Diversity

Analogous Environments Across the Tropics Have Similar Levels of Tree Species Alpha Diversity

Shu-Mei Xiao[1,4], Jonathan M. Adams[2*], Shu-Feng Li[1*], Ferry Slik[3], Daniel M. Griffith[5], Adriano Quaresma[6,7], 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 Harrison[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], Francisco Mora Ardila[36], Giselda Durigan[37], Gopal Shukla[38], Guillermo Ibarra-Manríquez[36], Hidetoshi Nagamasu[39], Ida Theilade[40], Irie Casimir Zo-Bi[41], Isau Huamantupa-Chuquimaco[42,43], J. Orlando Rangel-Ch.[44], James Grogan[45], Javid Ahmad Dar[46,47], Jochen Schöngart[48], John Herbohn[49], John R. Poulsen[50], John N. Williams[51], Jon Lovett[52], Jose Don De Alban[53,54], José Rafael Lozada[55], José Roberto Rodrigues Pinto[56], Juan Ernesto Guevara-Andino[57], Jurgi Cristóbal-Azkarate[58], Jürgen Homeier[59], Katrin Böhning-Gaese[60], , Khalid Rehman Hakeem[61,62], Kenneth Feeley[63], Kyle W. Tomlinson[64], Ladan Rasingam[65], Layon Oreste Demarchi[48], Laumonier Yves[66], Luciana F. Alves[67], Luis Torres Montenegro[68], Manichanh Satdichanh[69], Manuel J. Macía[70,71], Marcelo Tabarelli[72], Marcio Seiji Suganuma[73], Marcos Antonio Ríos Paredes[68,74], Maria Teresa Fernandez Piedade[48], Mark Schulze[75], Mattheas van de Bult[76], Meredith L. Bastian[77,78], Mohammed Latif Khan[79], Mohammad Shah Hussain[8], Michael Kessler[80], Michael J. Lawes[81], Miguel A. Munguía-Rosas[82], Narayanaswamy Parthasarathy[16], Naret Seuaturian[83], Naveen Babu Kanda[16,19, 84] , Naveenkumar Jagadeesan[31], Nigel Pitman[85], Onrizal Onrizal[86], Ole R. Vetaas[87], Pantaleo Munishi[88], Phourin Chhang[89], Polyanna da Conceição Bispo[90], Rahmad Zakaria[91], Rahayu Sukmaria Sukri[92], Rajkumar Muthu[93], Rama Chandra Prasad[94], Ramachandran, V.S.[95], Rhett D. Harrison[96], Rizza Karen Veridiano[97,98], Robert Steinmetz[83], Robin L Chazdon[99], Roven Tumaneng[98,100], Samir Gonçalves Rolim[22], SC Garkoti[101], Selene B áez[102], Serge Wich[103], Sharif A. Mukul[104,105], Shijo Joseph[106], Simone Aparecida Vieira[107], S. Muthuramkumar[108], Somaiah Sundarapandian[16], Sumit Chakravarty[109], Subashree Kothandaraman[46,47], Tânia Wendt[110], Thiago Metzker[111,112], Timothy Whitfeld[113], Tao Su[1,114], Van Andel Tinde R[115], Verbeeck Hans[116], Víctor Arroyo-Rodrí guez[36,117], Wanlop Chutipong[118], William F. Laurance[119], Yrma Andreina Carrero[120], Zhe-Kun Zhou[1]

[*]To whom correspondence may be addressed. Email: geog.ecol@gmail.com, jonadams@nju.edu.cn and lisf@xtbg.org.cn

(Both corresponding authors contributed equally to this manuscript)

[1]State Key Laboratory of Plant Diversity and Specialty Crops, Xishuangbanna Tropical Botanical Garden, Chinese Academy of Sciences, Mengla, Yunnan, 666303, China; [2]Nanjing University, Nanjing 210023, China; [3]Environmental and Life Sciences, Faculty of Science, University Brunei Darussalam, Jalan Tunku Link, Gadong, BE1410, Brunei Darussalam; [4]University of Chinese Academy of Sciences, Beijing 100049, China; [5]Departamento de Ciencias Biológicas y Agropecuarias, EcoSs Lab, Universidad Técnica Particular de Loja, Loja, 110150, Ecuador; [6]Grupo de Pesquisa Ecologia, Monitoramento e uso Sustentável de Áreas Úmidas (MAUA), Instituto Nacional de Pesquisas da Amazônia (INPA), Manaus, 69067-375, Brazil; [7]Department of Wetland Ecology, Institute of Geography and Geoecology, Karlsruhe Institute of Technology (KIT), Karlsruhe, 76131, Germany; [8]Biodiversity Parks Programme, CEMDE, Department of Environmental Studies, University of Delhi, Delhi, 110007, India; [9]Research Center for Plant Conservation, Botanic Gardens, and Forestry - National Research and Innovation Agency (BRIN), Bogor, 16122, Indonesia; [10]Facultad del Medio Ambiente y Recursos Naturales, Universidad Distrital Francisco José de Caldas, 110231, Bogotá, Colombia; [11]Forest Research Institute, University of the Sunshine Coast, Sippy Downs, QLD 4556, Australia; [12]Reforest Africa, Mang'ula, Kilombero, PO Box 5, Tanzania; [13]Flamingo Land

Ltd., Kirby Misperton, North Yorkshire, YO17 6UX, UK; [14]Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, Michigan, 48109, USA; [15]Oikobit LLC, www.oikobit.com, 2105 Vista Oeste St NW, Albuquerque, New Mexico, 87120, USA; [16]Department of Ecology and Environmental Sciences, School of Life Sciences, Pondicherry University, Puducherry - 605014, India; [17]School of Biological Sciences (SBS), Universiti Sains Malaysia, Penang, 11800, Malaysia; [18]Centre for Marine and Coastal Studies (CEMACS), Universiti Sains Malaysia, Penang, 11800, Malaysia; [19]Department of Ecology, French Institute of Pondicherry, Puducherry - 605001, India; [20]CIRAD, Bat D Bur 104, Campus de Baillarguet, Montpellier Cedex 5, 34398, France; [21]Department of Plant Biology, Institute of Biology, University of Campinas, UNICAMP, Campinas, 13083-862, Brazil; [22]Laboratório de Dendrologia e Silvicultura Tropical, Centro de Formação em Ciências Agroflorestais, Universidade Federal do sul da Bahia, Itabuna, 45613-204, Brazil; [23]Royal Botanic Garden Edinburgh, EH3 5LR, UK; [24]Harold L. Lyon Arboretum, University of Hawai'i at Mānoa, Honolulu, HI, 96822, USA; [25]Forest Ecology and Forest Management Group, Wageningen University & Research, PO Box 47, Wageningen, 6700 AA, The Netherlands; [26]International Institute for Sustainability, IIS-Rio, Rio de Janeiro, 22460-320, Brazil; [27]Programa de P ós-graduação em Ecologia e Conservação, Universidade do Estado de Mato Grosso, Nova Xavantina, 78690-000, Brazil; [28]Department of Biology, Universitas Bangka Belitung, Kampus Terpadu UBB Balunijuk, Kec. Merawang, Kab. Bangka, Provinsi Kepulauan Bangka Belitung 33172, Indonesia; [29]Herbarium Bangka Belitungense, Kampus Terpadu UBB Balunijuk, Kec. Merawang, Kab. Bangka, Provinsi Kepulauan Bangka Belitung 33172, Indonesia; [30]Departamento of Ecology and Conservation, Universidade Federal de Lavras, Lavras, MG, 37200-900, Brazil; [31]Viikki Tropical Resources Institute, Department of Forest Sciences, University of Helsinki, Latokartanonkaari 7, Helsinki, 00790, Finland; [32]Helsinki Institute of Sustainability Science (HELSUS), Yliopistonkatu 3, Helsinki, 00100, Finland; [33]Institute of Tropical Forestry and Forest Products (INTROP), Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia; [34]Federal Institute of Espirito Santo, Cariacica Campus, Espirito Santo, 29150-410, Brazil; [35]National Institute of the Atlantic Forest, Santa Teresa, Espirito Santo, 29650-000, Brazil; [36]Instituto de Investigaciones en Ecosistemas y Sustentabilidad, Universidad Nacional Autónoma de México. Morelia 58190, Michoacán, México; [37]Floresta Estadual de Assis, Instituto de Pesquisas Ambientais. PO Box 104, 19800-000, Assis, Brazil; [38]Department of Forestry, North Eastern Hill University, Tura Campus, Tura-794002, Meghalaya, India; [39]The Kyoto University Museum, Kyoto University, Kyoto 606-8501, Japan; [40]Department of Food and Resource Economics, University of Copenhagen, Rolighedsvej 23, 1958 Frb. C..; [41]Institute National Polytechnique Félix Houphouët-Boigny, INP-HB, BP 1093; [42]Herbario Alwyn Gentry (HAG), Departamento Académico de Ciencias B ásicas. Universidad Nacional Amazónica de Madre de Dios (UNAMAD), Av. Jorge Chávez 1160, Madre de Dios, 17001, Perú; [43]Centro Ecológico INKAMAZONIA, Valle de Kosñipata, vía Cusco - Parque Nacional del Manú. Cusco, Perú; [44]Instituto de Ciencias Naturales, Universidad Nacional de Colombia.Bogotá D.C.; [45]Botanic Garden of Smith College, Northampton, MA USA 01063; [46]Department of Environmental Science and Engineering, School of Engineering and Sciences, SRM University-AP, Andhra Pradesh-522240, India; [47]Centre for Geospatial Technology, SRM University-AP, Andhra Pradesh-522240, India; [48]Ecology, Monitoring and Sustainable Use of Wetlands (MAUA), Instituto Nacional de Pesquisas da Amazônia (INPA), Av. André Araújo, 2936, Aleixo, 69060-001, Manaus, Brazil; [49]Forest Research Institute, University of the Sunshine Coast, Sippy Downs, 4556, Queensland, Australia; [50]The Nature Conservancy 2424 Spruce St., Boulder, CO, 80302, USA; [51]University of California, Davis. Dept. Env. Sci. & Policy, One Shields Ave., Davis, CA 95616 USA; [52]School of Geography, University of Leeds, Leeds LS2 9JT, UK; [53]ClearWind Pte Ltd, Singapore, Singapore 068894 Fauna; [54]Flora International - Philippines Programme, Cambridge, CB2 3QZ United Kingdom; [55]Instituto de Investigaciones para el Desarrollo Forestal (INDEFOR), Universidad de Los Andes, Vía Chorros de Milla, Mérida 5101, Venezuela; [56]Departamento de Engenharia Florestal, Centro de Referência em Conservação da Natureza e Recuperação de Áreas Degradadas, Universidade de Brasília, Brasília 70910-900, Brazil; [57]Grupo de Investigación en Ecología y Evolución en los Tró

picos-EETrop, Universidad de las Américas, Quito 170125, Ecuador; [58]Department of Basic Psychological Processes and Development, University of the Basque Country UPV/EHU, 20018 Donostia, Spain; [59]Faculty of Resource Management, HAWK University of Applied Sciences and Arts, Göttingen, 37085, Germany; [60]Senckenberg Biodiversity and Climate Research Centre, Frankfurt am Main, Germany and Goethe University, Frankfurt am Main, 60323, Germany; [61]Department of Biological Sciences, Faculty of Science, King Abdulaziz University, Jeddah-21589, Saudi Arabia; [62]Princess Dr. Najla Bint Saud Al-Saud Center for Excellence Research in Biotechnology, King Abdulaziz University, Jeddah, 21589, Saudi Arabia; [63]Biology Department, University of Miami, Coral Gables, Florida, 33146, USA; [64]Center for Integrative Conservation & Yunnan Key Laboratory for Conservation of Tropical Rainforests and Asian Elephants, Xishuangbanna Tropical Botanical Garden, Mengla 666303, Yunnan, China; [65]Botanical Survey of India, Deccan Regional Centre, Hyderabad, Telangana, 500095, India; [66]Centre for International Forestry Research, Bogor, 16115, Indonesia; [67]Institute of the Environment and Sustainability, University of California Los Angeles, Los Angeles, California, 90095, USA; [68]Herbario Herrerense, Instituto de Investigaciones de la Amazonía Peruana, Iquitos, 16002, Peru; [69]School of Life Sciences, University of Hawai'i at Mānoa. 3190 Maile Way, St. John 101, HI 96822, Honolulu; [70]Departamento de Biología, Área de Botá nica, Universidad Autónoma de Madrid, 28049 Madrid, Spain; [71]Centro de Investigación en Biodiversidad y Cambio Global (CIBC-UAM), Universidad Autónoma de Madrid, 28049 Madrid, Spain; [72]Botany Department, Universidade Federal de Pernambuco, 50670-901, Brazil; [73]Colégio de Aplicação, Universidade Federal de Santa Catarina, Florian ópolis, 88040-900, Brazil; [74]Keller Science Action Center, The Field Museum, Chicago, IL 60605-2496; [75]H.J. Andrews Experimental Forest, Blue River, Oregon, 97413, USA; [76]Doi Tung Development Project, Social Development Department, Chiang Rai, 57240, Thailand; [77]Proceedings of the National Academy of Sciences, 500 Fifth St NW, Washington, DC 20001 USA; [78]Department of Evolutionary Anthropology, Duke University, Durham, NC 27708 USA; [79]Department of Botany, Dr. Harisingh Gour Vishwavidyalaya (A Central University), Sagar - 470003, Madhya Pradesh, India; [80]Systematic and Evolutionary Botany, University of Zurich, Zollikerstrasse 107, 8008 Zurich, Switzerland; [81]School of Life Sciences, University of KwaZulu-Natal, Scottsville 3209, South Africa; [82]Laboratorio de Ecología Terrestre, Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional (Cinvestav). Mérida 97310, Yucatán, Mexico; [83]World Wildlife Fund Thailand, Bangkok, 10400, Thailand; [84]National Biobank of Thailand, National Science and Technology Development Agency, Pathum Thani, 12120, Thailand; [85]Field Museum of Natural History 1400 S. DuSable Lake Shore Dr. Chicago, IL 60605 USA; [86]Faculty of Forestry, Universitas Sumatera Utara, Medan, 20155, Indonesia; [87]Department of Geography, University of Bergen, PO Box 7802, N - 5020 Bergen, Norway; [88]Department of Ecosystems and Conservation, Sokoine University of Agriculture, PO Box 3010 Morogoro Tanzania; [89]Forest and Wildlife Research and Development Institute, Cambodia; [90]Department of Geography, School of Environment Education and Development, University of Manchester, Oxford Road, M13 9PL, Manchester, UK; [91]School of Biological Sciences, Universiti Sains Malaysia, Pulau Pinang, 11800, Malaysia; [92]Institute for Biodiversity and Environmental Research, Universiti Brunei Darussalam, Jalan Tungku Link, BE 1410, Brunei Darussalam; [93]Tropical Forest Research Institute, Jabalpur, Madhya Pradesh, India – 482021; [94]Lab for Spatial Informatics, International Institute of Information Technology, Hyderabad, Gachibowli, Telangana, India – 500032; [95]Centre for Environmental Studies, Amrita School of Engineering, Coimbatore, Amrita Vishwa Vidyapeetham, 641112, India; [96]CIFOR-ICRAF St Eugene Office Park, Lake Road, 10101, Lusaka Zambia; [97]FORLIANCE GmbH, 53119 Bonn, Germany; [98]Fauna & Flora International - Philippines Programme, Cambridge, CB2 3QZ, United Kingdom; [99]Forest Research Institute, University of the Sunshine Coast, 90 Sippy Downs Dr, Sippy Downs QLD, 4556; [100]Department of Science and Technology - Philippine Council for Industry, Energy and Emerging Technology Research and Development, Taguig City, 1631 Philippines; [101]School of Environmental Sciences, Jawaharlal Nehru University, New Delhi- 110067, India; [102]Department of Biology, Faculty of Sciences, National Polytechnic School of Ecuador, 170525, Quito-Ecuador;

[103]School of Biological and Environmental Sciences, Liverpool John Moores University, James Parsons Building, Byrom street, L33AF, Liverpool, UK; [104]Department of Environment and Development Studies, United International University, Dhaka 1212, Bangladesh; [105]Tropical Forests and People Research Centre, University of the Sunshine Coast, Maroochydore DC, Queensland, 4558, Australia; [106]Department of Climate Variability and Aquatic Ecosystems, Faculty of Ocean Science and Technology, Kerala University of Fisheries and Ocean Studies, Kochi, Kerala - 692508, India; [107]Environmental Studies and Research Center, Universidade Estadual de Campinas, UNICAMP, Campinas, 13083-970, Brazil; [108]Research Centre in Botany, VHNSN College, Virudhunagar - 626001, India; [109]Department of Forestry, Uttar Banga Krishi Viswavidyalaya, Pundibari-736165, Cooch Behar, West Bengal, India; [110]Department of Botany, Institute of Biology, Federal University of Rio de Janeiro, Rio de Janeiro, 21941-590, Brazil; [111]IBAM - Instituto Bem Ambiental, Belo Horizonte, 30130-090, Brazil; [112]Myr Projetos Sustentáveis (Grupo Myr), Belo Horizonte, 30130-009, Brazil; [113]Bell Museum, University of Minnesota, St. Paul, Minnesota, 55108, USA; [114]State Key Laboratory of Oil and Gas Reservoir Geology and Exploitation & Institute of Sedimentary Geology, Chengdu University of Technology, Chengdu 610059, China; [115]Naturalis Biodiversity Center, PO Box 9517, 2300 RA Leiden, the Netherlands; [116]CAVElab -Computational & Applied Vegetation Ecology, Department of Environment, Faculty of Bioscience Engineering, Ghent University, Coupure Links 653, 9000 Gent, Belgium; [117]Escuela Nacional de Estudios Superiores, Universidad Nacional Autónoma de México, Merida 97357, Yucatán, Mexico; [118]Conservation Ecology Program, King Mongkut's University of Technology Thonburi, Bangkok, 10150, Thailand; [119]Centre for Tropical Environmental and Sustainability Science, and College of Science and Engineering, James Cook University, Cairns, Queensland 4878, Australia; [120]Programa de Maestría en Manejo de Bosques, Universidad de Los Andes , Vía Chorros de Milla, Mérida 5101, Venezuela.

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 (RF) model was trained using a network of 429 one-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 around 86% of variation in richness. Despite differences in regional species pools and 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 regional species assemblages. This pantropical convergence in the richness-environment relationship poses a challenge for ecology to explain.

Keywords: rainforest, tree richness, modelling, 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 using least-squares regression to reveal a general relationship between primary productivity and species richness. Other studies of woody plant species richness, reviewed by [6], have concentrated on analyzing 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, less 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: a) Local-scale tree alpha diversity across the global tropics can be consistently predicted by contemporary environmental variables, such that analogous environmental conditions will support similar levels of richness, irrespective of deep historical biogeographic divergences; b) 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 on the local scale compared to 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.

198 Comparing 2046 tree plots across Amazonia, Ter Steege et al. [8] demonstrated that a
199 combined influence of climate and soil factors explained local tree richness and
200 community composition. In a comparison of forest plots from South America and
201 Africa, Parmentier et al. [9] showed lower plot-level richness in Africa under similar
202 warm and moist climatic conditions. In global surveys of forest plots of varying sizes,
203 Keil and Chase [10] and Chu et al. [11] found that the drivers of variation in diversity
204 differ with plot size and spatial distance amongst samples, which are important
205 sampling issues in ecology generally [12]. However, no studies thus far have addressed
206 local tree species richness using a large number of plots of standardized size across the
207 world's tropics, which is necessary for continental-scale comparisons. The increasing
208 availability and integration of forest plot data now make such comparisons possible
209 [13].
210 Two other important advances in environmental data science facilitate hypothesis
211 testing about the patterns and processes driving tropical tree diversity. First, the
212 availability of interpolated climate, soil and other environmental parameters has grown
213 immensely in recent years [14–18]. Second, machine learning models enable analysis
214 of the simultaneous effect of multiple factors on community structure and diversity
215 across large scales [19,20]. A particularly robust machine learning method is random
216 forest (RF) modeling which, when combined with spatial regression, enables analysis
217 and prediction of spatially structured data where observations exhibit autocorrelation
218 [21,22]. These models extend the capabilities of traditional RF to handle spatial data
219 more effectively, allowing for spatial autocorrelation and other geographic phenomena
220 to be incorporated into vegetation modeling.
221 Here we used the Pantropical Forests Network (PFN) of tree inventory plots assembled
222 by Slik et al. [23] in combination with publicly available environmental data surfaces
223 [14–18] to determine which environmental variables explain plot-level tree α diversity
224 across the global tropics. We applied a spatial RF and negative binomial generalized
225 linear model to detect and compare the strength of empirical links between
226 environmental variables and tree species richness. The number of species per unit area
227 (e.g. Gentry [1]) - the classic measure of forest tree richness - has its weaknesses given

that stem density varies by both biogeographical region and along climate gradients. To compare species richness while accounting for variations in sampling effort and completeness, we employed rarefaction techniques, sample-coverage-based approaches based on abundance data (e.g., [24,25]). And Fisher's α [26], and the 'classical' method of 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: (1) Which environmental predictors explain most of the variation in local tree species richness across the tropics? (2) How do the environmental factors associated with tree species richness differ between broad versus local scales of sample plot spacing? (3) 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 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, 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), precipitation of the driest month (bio14), and annual range of monthly relative humidity (hurs_range), along with soil pH, isothermality (bio03), NPP, and annual range (rsds_range) and mean (rsds_mean) of monthly surface downwelling shortwave flux in air (Fig. 1, Table S2). PC2 was predominantly

comprised of soil-related variables, including total nitrogen, volumetric water content (wv0010), and SOC, as well as topography-related variables including tangential curvature (tcurv), ruggedness, and vector ruggedness measurement (vrm) (Fig. 1, 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 to Asia and Africa (Fig. 1). Specifically, Asia demonstrated the highest precipitation during the driest month and lower precipitation seasonality, while Africa showed the least 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 alpha form standardized sample of stems. By analyzing non-spatial and spatial RF based on observed richness, Fisher's alpha and richness from rarefaction data returned high 'out-of-bag' R^2 (OBB) 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, Table S1).

Training the RF for each of the three richness measures revealed that predicted tree species richness at the 1-ha scale was highly heterogeneous across the tropics (Fig. 4A, Fig. S6A, Fig. S7A), with 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 was found across most of tropical Africa, eastern and southern Amazonia, and continental Southeast Asia.

Analysis using non-spatial RF revealed that precipitation of the driest month (bio14) was identified as the most important single variable for predicting tree richness, followed by soil pH (phh2o) (Fig. 4B). The interaction between precipitation of the driest month and 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 Supplementary Methods 1), showed higher importance of moisture-related variables, including annual range of monthly relative humidity (hurs_range), precipitation seasonality (bio15), mean monthly potential

evapotranspiration (pet_penman_mean), and the mean monthly climate moisture index (cmi_mean) (Fig. 4C). In contrast, the importance of soil-related variables, such as pH and silt, remained consistent, while the interaction between precipitation of the driest month and silt (bio14 x pca..silt) decreased in importance. An additional analysis in which non-spatial and spatial RF was 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 like pH, soil organic carbon (soc), and silt (Fig. 4D, Fig. S8B, D). This reduction was particularly pronounced with 100 and 200-km thinning, where the importance of soil pH and silt decreased markedly. In contrast, variables associated with moisture (e.g., annual range of monthly relative humidity (hurs_range), mean monthly climate moisture index (cmi_mean)), growing season (e.g., net primary productivity [NPP]), solar radiation (e.g., mean monthly surface downwelling shortwave flux [rsds_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: precipitation of the driest month (bio14), soil pH (phh2o), annual range of monthly relative humidity (hurs_range), precipitation seasonality (bio15), silt proportion (silt), and mean monthly potential evapotranspiration (pet_penman_mean) (Fig. S9). In all three regions, tree richness was highest in areas with abundant moisture in the driest month (bio14; Fig. S9A). Conversely, richness was low in areas with high soil pH and silt fraction, except in Asia, where richness increased with silt (Fig. S9B, E). High tree richness was also associated with low precipitation seasonality, annual range of monthly relative humidity, and mean monthly potential evapotranspiration, peaking where these were approximately 25–35 mm, 5−6%, and 110–115 mm/month, respectively (Fig. S9C, D, F). Notably, the Americas displayed large variation in precipitation of the driest month, soil pH and silt, whereas Asia exhibited substantial variation in humidity range, precipitation seasonality, and monthly potential evapotranspiration. Africa showed low variation in these parameters compared to the other two regions.

The contribution of the six dominant environmental variables to local tree species

richness varied within and across regions. In areas with high precipitation, low seasonality in rainfall, and low soil pH (Fig. S10–11), the predictive contribution of these climatic variables to species richness within the RF model was reduced (Fig. S12–13). For example, in areas with low soil pH (around 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.

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 the Americas (Tables S6–7). Despite those differences, the direction and significance of the effects of important precipitation (e.g., precipitation of the driest month [bio14], precipitation seasonality [bio15], mean monthly vapor pressure deficit [vpd_mean], annual range of monthly relative humidity [hurs_range]), temperature (mean monthly minimum air temperature of the coldest month [bio06]) and soil variables (e.g., pH, silt) were similar between the two approaches (Tables S6–7). Variables showing a significant interaction between Africa and the Americas included cmi_mean, hurs_range, vpd_mean, clay proportion, soil pH, and tcurv (Tables S6–7). Both the glm.nb and RF models showed consistent results, indicating that variables like bio14, pH, hurs_range, silt, and bio15 are important factors in controlling alpha diversity. While RF excels in predictive power ($R^2 = 0.86$), capturing complex, non-linear relationships, the glm.nb model ($R^2 = 0.69$) also explains a large portion of the richness. This demonstrates glm.nb's ability to capture substantial data variability despite its lower R^2. Crucially, glm.nb shows significant regional differences between 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 non-linear 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, co-limitation that means no single type of environmental factor dominates, local tree species richness was largely explained by the same three categories: co-limitation, precipitation and soil (Fig. 5). Referring to areas where 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 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 richness variation in the three regions. Compare with other regions, along the longitudinal gradient, growing season and solar radiation were important in Africa, while 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 and glm.nb model showed that variation in tropical tree species richness was consistently predictable by 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 observed raw richness and predicted richness showing similar results in RF (Table S5), and relatively lower in 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 Random Forest 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 inclusion of additional forest plots (taking in a greater range of local environments or sub-regions), the sample size as it already stands is very large, and in diverse settings, so we regard these inferences be robust and unlikely to change substantially with addition of more samples.

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 on 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 (Fig. 4, Fig. 5) were those reflecting aspects of moisture supply to vegetation, with variation in temperature playing a less important role. The estimated net primary productivity (NPP) itself is high on the list of predictors at large scales (50

km, 100 km and 200 km, Fig. 4D, Fig. S8). In general, the climate factors that emerge as important 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 by 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.

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 coexistence of high species richness.

Certain internal patterns of influence on species richness within each tropical region are discernable (Fig. 5, Fig. S10–13). 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 variation in 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 where resources are relatively balanced, favoring the survival of various species and thus making these areas rich in species richness. Given the

relatively limited attention given to the role of soil parameters in affecting patterns of tropical tree diversity, soil pH surprisingly emerges as influential in the RF analysis, especially at the local scale (Fig. 4B–C, Fig. S8). The role of soil pH has long been known as an influence on more local scale patterns in tropical forest composition, at least [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]. Species composition of vegetation can itself influence 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 speculative.

While 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 (Fig. S9B, Fig. S10B). Empirically, this supports previous results from studies in other biomes that plant species richness peaks in low soil pH, around pH 4 [49], and declines as 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 trees as the primary driver. For instance, the Janzen-Connell hypothesis suggests that diversity levels in tropical forests are controlled by the intensity of attack by insect herbivores and pathogens [2,52–54] – with constantly warm and moist conditions favoring specialization of insect pest or pathogen populations, and 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 soil pH on herbivory – perhaps with respect to the extent of the species pools on different soil types and perhaps in affecting nutrient or secondary compound content of plants, or the growth rate of fresh edible tissues – might also be involved in the observed relationship to soil factors.

Convergence in richness amongst biogeographical 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 is 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, Fig. S14–16). The spatial RF model demonstrates that analogous environments support comparable richness levels, irrespective of regional species pools (Fig. S17). This contrasts with 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, Fig. S14–16). 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 correlates significantly with species richness independently of climate. When RF 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, SE Asia (and Wallacea/Oceania), and tropical Africa [38,56,58]. In each region, the composition and distribution of flora have been influenced by continental collisions, mountain building, and the cooling and drying of climate in the Cenozoic [30,56,57,59,60]. For instance, tropical Africa lost a large proportion of its previous diversity during the late Cenozoic/Quaternary due to drying of climate, especially during glacial episodes [38,56,57], and now has a restricted total richness and phylogenetic diversity of trees on 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 enclave 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 [61,62].

Despite all of these different histories and potential trajectories, brought about by climatic and tectonic history, and regional scale diversifications, it is striking that from the perspective of the 1-ha sample scale all the regions we distinguish here adhere closely to the same pantropical pattern of richness, in relation to present-day environmental conditions (Fig. 2, Fig. S14–16).

The very close correspondence in tree species richness between regions, despite all of the historical legacy factors that could potentially cause divergence in diversity, implies the existence of precise control by ‘governance’ factors in the forest community that tend to cause richness to settle at a particular level. Many potential mechanisms have been put forward to explain how 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 local scale richness, when there are so many factors that would be expected to cause 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 progressive buildup of diversity by diversification or migration, and its destruction by extinction [63]. According to such mechanisms, there is no 'lid' on 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 coexist 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 – for example the number of discrete niches available due to the heterogeneity of microenvironments[64], the narrowness of specialized niches that is possible in a given environment (affecting the opportunities for slotting in extra species) [65], or the degree of overlap in tree species niches that can occur before competitive exclusion begins to reduce diversity [66,67]. In our opinion, the results of this study support a predominance of equilibrium or niche-based mechanisms, since the hectare scale species richness is so strongly convergent between different parts of the world. If the vagaries of diversification and extinction were more important in affecting richness, 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 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 amongst 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

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

555 with tree species richness patterns across and between tropical regions. Pantropical
556 regions exhibit high species richness, primarily due to the intricate interplay of co-
557 limitation that create stable and balanced conditions. Although regional species pools
558 differ, analogous environmental conditions yield similar local richness patterns. A
559 multifactorial interplay of evolutionary and ecological mechanisms is presumably at
560 work controlling this consistency of species richness, preventing regional divergence.
561 The observed variation in tropical species richness must be seen as the outcome of
562 a range of different factors acting simultaneously, sometimes in parallel and
563 sometimes in opposition to one another. It is possible that these multiple factors
564 each act through a range of different ecological mechanisms, requiring separate
565 elucidation. These findings decisively strengthen the case for niche theory, revealing
566 its remarkable predictive power. They provide compelling evidence that niche
567 segregation is a pervasive and dominant force, structuring communities from local
568 patches to broad regional landscapes, even in the presence of diverse regional species
569 assemblages.
570 Whilst identification of the true underlying mechanisms involved remains a
571 fundamental challenge for ecology, this study contributes to the ongoing challenge in
572 ecology to identify and understand the controls on biological diversity.
573 In this paper, we had originally hypothesized that local-scale tree alpha diversity
574 across the global tropics can be consistently predicted by contemporary
575 environmental variables, such that analogous environmental conditions will
576 support similar levels of richness, irrespective of deep historical biogeographic
577 divergences. This hypothesis has survived its test, with striking predictability of species
578 richness as sampled at the hectare scale. We also hypothesized that variation in richness
579 on more localized scales between samples would be governed by a distinct set of
580 influences. Despite some subtle scale-related differences, this was essentially disproven,
581 with similar sets of environmental factors governing throughout.
582 It is necessary to keep in mind, however, that our study is confined to data obtained at
583 the 1 ha sampling scale, and that other patterns may emerge at other local sampling
584 scales [10], or for other life forms (e.g. lianas) – either in sub-hectare or larger plots, or

585 at the level of beta-diversity turnover rates. This awaits other studies, and the additional
586 information that comes from these will shed more light on the mechanisms at work
587 behind tropical tree diversity patterns. The findings of this study are derived from
588 current sample surveys, and the uneven distribution of these samples may affect our
589 results. We advocate for an increase in sample surveys in tropical regions and the
590 development of a more comprehensive tropical sample database in future scientific
591 research.
592 Intriguingly, other aspects of community structuring of tropical forests may be found to
593 show striking convergence patterns across the tropics. Cooper et al. [68] have recently
594 revealed strong convergence in relative abundance data – rather than richness as is
595 shown here - in forest plots throughout the tropics. There is a need for further careful
596 comparisons of the structure and functioning of tropical forests across different regions,
597 to understand how closely they have maintained their similarities, and to better
598 understand the driving mechanisms behind the observed patterns.
599
600 **Materials and Methods**
601 **Tree data and richness**
602 Tree inventory data were assembled from the Pantropical Forests Network (PFN) of
603 old-growth (not recently logged or cleared) closed canopy forest plots from across the
604 global tropics, including tropical dry forest and its various transitional forms to tropical
605 rainforest [23]. All trees, defined as free-standing woody individuals (including palms),
606 with a diameter at breast height (1.3 m) ≥10 cm was measured and identified in each
607 plot (Data S1). If species names could not be determined, plot-specific morpho-species
608 were recorded, with closest taxonomic assignment. All morpho-types included here
609 were identified to the Linnean species, genus or at least family level. Unknowns at
610 Linnean family level were not included.
611 The numbers of individuals and species were obtained from a total of 429 1-ha plots
612 located in the Americas (197 plots), Africa (150 plots) and Asia (82 plots) (Fig. 2, Fig.
613 S3–S4, Data S2). The Asian region included Wallacea and Australasia based on shared
614 floristic affinity [23]. Geographical boundaries of the tropics were defined based on the

map of [69].

In addition to observed species richness, we analyzed the richness in each 1-ha plot estimated from sample-based rarefaction and Fisher's α (see Supplementary 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 [25,70]. This technique allows for 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 [71] to estimate species richness based on sample coverage, which accounts for the completeness of sampling by estimating how well a community has been sampled. 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 regions. The rarefaction curves were plotted to visualize the richness of species across coverage sampling efforts, indicate sample coverage values for all localities are quite high, most of them are larger than 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 where species follow a log-series pattern with high proportions of rare species and for detecting the influence of abundance distributions on species diversity [26,72]. Compared to other diversity measures, these two indices showed the strongest correlation with observed richness (Table S4, Fig. S2, see Supplementary 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 (NDVI), and climatic background information [15]. Topographic data with a spatial resolution of approximately 1 km were downloaded from EarthEnv (http://www.earthenv.org/topography). Important environmental variables were selected using the Boruta algorithm from the *Boruta* package [73] in R (Data S6). Multicollinearity was addressed by assessing correlation and variance inflation factors, resulting in a refined set of 24 environmental predictors for RF analysis (Table S1).

Spatial random forest

To predict species richness based on multiple environmental variables, we applied a random forest (RF) combined with spatial regression using the package *spatialRF* in R [22], which can be applied on regular or irregular data [74]. This package enhances traditional RF techniques by accounting for spatial autocorrelation where observations are not independent but rather show geographic relationships. Spatial autocorrelation based on Moran's I index was explicitly taken into account to improve 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 using the "thin" function from the *spThin* package [75], 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 results from the full dataset (see Supplementary methods S2).

Statistical analysis

We used linear regression to examine the relationship between species richness predicted by non-spatial and spatial RF and that estimated from sample-based

675 rarefaction, Fisher's α and observed richness. Principal component analysis (PCA) was
676 performed to explore variation among the 24 environmental variables in the three major
677 tropical regions: the Americas, Africa and Asia. We used the *FactoMineR* package [76]
678 in R to implement PCA and the *factoextra* package [77] to generate a biplot that enabled
679 visualization of the results.
680 Spatial Random Forest (RF) excels at capturing complex, non-linear relationships
681 and identifying important predictors, primarily for its predictive power [78]. As a
682 complementary approach, the Generalized Linear Model with Negative Binomial
683 distribution (glm.nb) was used. This allowed us to validate RF results by
684 statistically testing broad-scale factors such as continental differences and
685 examining linear relationships, combining data-driven prediction with statistical
686 inference. To quantify the effects of the 24 environmental variables on species richness
687 across the three tropical regions, we applied a negative binomial generalized linear
688 model (glm.nb) using the *MASS* package in R and compared the results with those from
689 spatial RF. Spatial RF excels in capturing complex, non-linear relationships between
690 species richness and environmental variables, but does not provide interpretable
691 coefficients that explain the direction and magnitude of those relationships [79]. As a
692 complementary approach, glm.nb was used to validate the robustness of the RF results
693 by enabling the identification and interpretation of regional differences and linear
694 relationships between environmental predictors and species richness. Both main effects
695 and interactions between environmental variables and region were examined with this
696 approach (see Supplementary methods S3).
697
698 **Sensitivity analysis**
699 To determine which class of environmental variables (e.g., precipitation, temperature,
700 soil, etc.) best explained variation in local tree species richness, we classified the 24
701 environmental variables into seven categories – temperature, precipitation, growing
702 season, Solar radiation, Soil, Topography and conducted a sensitivity analysis of the RF
703 model following the methods of Saltelli et al. [80] and Liang et al. [19]. All of the above
704 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 predicted species richness $Y_{all}(s)$:

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

where $f()$ represents the RF model, $X(s)$ represents the values of environmental variables, and s represents 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 tree species richness based on all 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 variables except those associated with temperature and $(X - E1)$ (s) represents the variables of the remaining 5 categories (E2-E6).

Step 3: We calculated the relative sensitivity of 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 $\geq 1/7$ was considered that which best explained the variation in tree richness for that area.

Step 5: In areas where relative sensitivities were less than 1/7 for all categories, we hypothesized that 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 where 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 regional variation in category importance, we calculated the relative sensitivity of each category as a percentage of that of all categories and plotted this along latitudinal and longitudinal gradients spanning the

tropics on a map.

SUPPLEMENTARY MATERIALS

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

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.

Author contributions

Shu-Mei Xiao conducted analysis, writing and interpretation. Shu-Feng Li conducted analysis, writing and interpretation. Jonathan M. Adams initiated and conceptualized the study, conducted interpretation, and led the writing. Ferry Slik coordinated data organization and assisted in interpretation. Kyle W. Tomlinson assisted in data interpretation. Tao Su and Zhe-Kun Zhou review and editing. PFN members D.M.G., 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.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.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.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., Y.A.C. contributed research data.

FUNDING

This work was supported by the National Key R&D Program of China (No. 2022YFF0800800), the National Natural Science Foundation of China (NSFC) (No. 42320104005; No. 42372033), the Young and Middle-Aged Academic and Technical Leaders of Yunnan (No. 202305AC160051), and the 14th Five-Year Plan of the Xishuangbanna Tropical Botanical Garden, Chinese Academy of

Sciences (No. E3ZKFF7B).

Conflict of interest statement. None declared.

References

[1] Gentry AH. Changes in Plant Community Diversity and Floristic Composition on Environmental and Geographical Gradients. Annals of the Missouri Botanical Garden 1988;75:1–34. https://doi.org/10.2307/2399464.

[2] Hill JL, Hill RA. Why are tropical rain forests so species rich? Classifying, reviewing and evaluating theories. Progress in Physical Geography: Earth and Environment 2001;25:326–54. https://doi.org/10.1177/030913330102500302.

[3] Hubbell SP. The unified neutral theory of biodiversity and biogeography Princeton 2001.

[4] Adams JS. Species richness: patterns in the diversity of life. Berlin ; New York : Chichester, UK: Springer ; in Association with Praxis; 2009.

[5] Houlahan JE, McKinney ST, Anderson TM, McGill BJ. The priority of prediction in ecological understanding. Oikos 2017;126:1–7. https://doi.org/10.1111/oik.03726.

[6] Hawkins BA, Albuquerque FS, Araújo MB, Beck J, Bini LM, Cabrero-Sañudo FJ, et al. A Global Evaluation of Metabolic Theory as an Explanation for Terrestrial Species Richness Gradients. Ecology 2007;88:1877–88. https://doi.org/10.1890/06-1444.1.

[7] Ricklefs RE, He F. Region effects influence local tree species diversity. Proceedings of the National Academy of Sciences 2016;113:674–9. https://doi.org/10.1073/pnas.1523683113.

[8] Ter Steege H, Pitman NCA, Do Amaral IL, De Souza Coelho L, De Almeida Matos FD, De Andrade Lima Filho D, et al. Mapping density, diversity and species-richness of the Amazon tree flora. Commun Biol 2023;6:1130. https://doi.org/10.1038/s42003-023-05514-6.

[9] Parmentier I, Malhi Y, Senterre B, Whittaker RJ, and ALONSO A, Balinga MPB, et al. The odd man out? Might climate explain the lower tree α-diversity of African rain forests relative to Amazonian rain forests? Journal of Ecology 2007;95:1058–71. https://doi.org/10.1111/j.1365-2745.2007.01273.x.

[10] Keil P, Chase JM. Global patterns and drivers of tree diversity integrated across a continuum of spatial grains. Nat Ecol Evol 2019;3:390–9. https://doi.org/10.1038/s41559-019-0799-0.

[11] Chu C, Lutz JA, Král K, Vrška T, Yin X, Myers JA, et al. Direct and indirect effects of climate on richness drive the latitudinal diversity gradient in forest trees. Ecology Letters 2019;22:245–55. https://doi.org/10.1111/ele.13175.

[12] Chase JM, McGill BJ, Thompson PL, Antão LH, Bates AE, Blowes SA, et al. Species richness change across spatial scales. Oikos 2019;128:1079–91. https://doi.org/10.1111/oik.05968.

[13] ForestPlots.net, Blundo C, Carilla J, Grau R, Malizia A, Malizia L, et al. Taking the pulse of Earth's tropical forests using networks of highly distributed plots. Biological Conservation 2021;260:108849. https://doi.org/10.1016/j.biocon.2020.108849.

[14] Karger DN, Conrad O, Böhner J, Kawohl T, Kreft H, Soria-Auza RW, et al. Climatologies at high resolution for the earth's land surface areas. Sci Data 2017;4:170122. https://doi.org/10.1038/sdata.2017.122.

[15] Hengl T, Jesus JM de, Heuvelink GBM, Gonzalez MR, Kilibarda M, Blagotić A, et al. SoilGrids250m: Global gridded soil information based on machine learning. PLOS ONE

2017;12:e0169748. https://doi.org/10.1371/journal.pone.0169748.
[16] Amatulli G, Domisch S, Tuanmu M-N, Parmentier B, Ranipeta A, Malczyk J, et al. A suite of global, cross-scale topographic variables for environmental and biodiversity modeling. Sci Data 2018;5:180040. https://doi.org/10/ggf2ks.
[17] Poggio L, de Sousa LM, Batjes NH, Heuvelink GBM, Kempen B, Ribeiro E, et al. SoilGrids 2.0: producing soil information for the globe with quantified spatial uncertainty. SOIL 2021;7:217–40. https://doi.org/10.5194/soil-7-217-2021.
[18] Brun P, Zimmermann NE, Hari C, Pellissier L, Karger DN. Global climate-related predictors at kilometer resolution for the past and future. Earth System Science Data 2022;14:5573–603. https://doi.org/10.5194/essd-14-5573-2022.
[19] Liang J, Gamarra JG, Picard N, Zhou M, Pijanowski B, Jacobs DF, et al. Co-limitation towards lower latitudes shapes global forest diversity gradients. Nature Ecology & Evolution 2022;6:1423–37. https://doi.org/10.1038/s41559-022-01831-x.
[20] Cai L, Kreft H, Taylor A, Denelle P, Schrader J, Essl F, et al. Global models and predictions of plant diversity based on advanced machine learning techniques. New Phytologist 2023;237:1432–45. https://doi.org/10.1111/nph.18533.
[21] Wright MN, Ziegler A. ranger: A Fast Implementation of Random Forests for High Dimensional Data in C++ and R. J Stat Soft 2017;77. https://doi.org/10.18637/jss.v077.i01.
[22] Benito BM. spatialRF: Easy Spatial Modeling with Random Forest 2022.
[23] Slik JWF, Arroyo-Rodriguez V, Aiba S-I and others. An estimate of the number of tropical tree species. Proceedings of the National Academy of Sciences 2015. https://doi.org/10/f7gq67.
[24] Gotelli NJ, Colwell RK. Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. Ecology Letters 2001;4:379–91.
[25] Chao A, Jost L. Coverage-based rarefaction and extrapolation: standardizing samples by completeness rather than size. Ecology 2012;93:2533–47. https://doi.org/10.1890/11-1952.1.
[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. The Journal of Animal Ecology 1943:42–58.
[27] Capinha C, Essl F, Seebens H, Pereira HM, Kühn I. Models of alien species richness show moderate predictive accuracy and poor transferability. NeoBiota 2018;38:77–96. https://doi.org/10.3897/neobiota.38.23518.
[28] Shafiullah AZ, Werner J, Kennedy E, Leso L, O'Brien B, Umstätter C. Machine Learning Based Prediction of Insufficient Herbage Allowance with Automated Feeding Behaviour and Activity Data. Sensors 2019;19:4479. https://doi.org/10.3390/s19204479.
[29] Slik JWF, Franklin J, Arroyo-Rodríguez V, Field R, Aguilar S, Aguirre N, et al. Phylogenetic classification of the world's tropical forests. Proceedings of the National Academy of Sciences 2018;115:1837–42. https://doi.org/10.1073/pnas.1714977115.
[30] Qian H, Kessler M, Zhang J, Jin Y, Soltis DE, Qian S, et al. Angiosperm phylogenetic diversity is lower in Africa than South America. Science Advances 2023;9:eadj1022. https://doi.org/10.1126/sciadv.adj1022.
[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. https://doi.org/10.1002/ece3.6419.

[32] Cooper DLM, Lewis SL, Sullivan MJP, Prado PI, ter Steege H, Barbier N, et al. Consistent patterns of common species across tropical tree communities. Nature 2024:1–10. https://doi.org/10.1038/s41586-023-06820-z.
[33] Whitmore TC. An introduction to tropical rain forests. 1990.
[34] Pires J, Prance G. The vegetation types of the Brazilian Amazon 1985.
[35] Currie DJ, Paquin V. Large-scale biogeographical patterns of species richness of trees. Nature 1987;329:326–7. https://doi.org/10.1038/329326a0.
[36] Saupe EE, Myers CE, Townsend Peterson A, Soberón J, Singarayer J, Valdes P, et al. Spatio-temporal climate change contributes to latitudinal diversity gradients. Nat Ecol Evol 2019;3:1419–29. https://doi.org/10.1038/s41559-019-0962-7.
[37] Coelho MTP, Barreto E, Rangel TF, Diniz-Filho JAF, Wüest RO, Bach W, et al. The geography of climate and the global patterns of species diversity. Nature 2023;622:537–44. https://doi.org/10.1038/s41586-023-06577-5.
[38] Morley RJ. Origin and evolution of tropical rain forests. 2000.
[39] Steenis van. Patio ludens and extinction of plants. 1978.
[40] Kleidon A, Adams J, Pavlick R, Reu B. Simulated geographic variations of plant species richness, evenness and abundance using climatic constraints on plant functional diversity. Environ Res Lett 2009;4:014007. https://doi.org/10.1088/1748-9326/4/1/014007.
[41] Reu B, Proulx R, Bohn K, Dyke JG, Kleidon A, Pavlick R, et al. The role of climate and plant functional trade-offs in shaping global biome and biodiversity patterns. Global Ecology and Biogeography 2011;20:570–81. https://doi.org/10.1111/j.1466-8238.2010.00621.x.
[42] Domingues TF, Meir P, Feldpausch TR, Saiz G, Veenendaal EM, Schrodt F, et al. Co-limitation of photosynthetic capacity by nitrogen and phosphorus in West Africa woodlands. Plant Cell & Environment 2010;33:959–80. https://doi.org/10.1111/j.1365-3040.2010.02119.x.
[43] Chapin FS, Bloom AJ, Field CB, Waring RH. Plant Responses to Multiple Environmental Factors. BioScience 1987;37:49–57. https://doi.org/10.2307/1310177.
[44] Grime JP. Plant Strategies, Vegetation Processes, and Ecosystem Properties. John Wiley & Sons; 2006.
[45] van Breemen N, Finzi AC. Plant-soil Interactions: Ecological Aspects and Evolutionary Implications. Biogeochemistry 1998;42:1–19. https://doi.org/10.1023/A:1005996009413.
[46] Kulmatiski A, Beard KH, Stevens JR, Cobbold SM. Plant–soil feedbacks: a meta-analytical review. Ecology Letters 2008;11:980–92. https://doi.org/10.1111/j.1461-0248.2008.01209.x.
[47] van der Putten WH, Bardgett RD, Bever JD, Bezemer TM, Casper BB, Fukami T, et al. Plant–soil feedbacks: the past, the present and future challenges. Journal of Ecology 2013;101:265–76. https://doi.org/10.1111/1365-2745.12054.
[48] Zhao S, 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, editors. Environment, Climate, Plant and Vegetation Growth, Cham: Springer Nature Switzerland; 2024, p. 229–64. https://doi.org/10.1007/978-3-031-69417-2_8.
[49] Azevedo LB, Van Zelm R, Hendriks AJ, Bobbink R, Huijbregts MAJ. Global assessment of the effects of terrestrial acidification on plant species richness. Environmental Pollution 2013;174:10–5. https://doi.org/10.1016/j.envpol.2012.11.001.
[50] Huston MA. Biological Diversity: The Coexistence of Species. Cambridge University Press;

1994.
[51] Phillips OL, Hall P, Gentry AH, Sawyer SA, Vásquez R. Dynamics and species richness of tropical rain forests. Proceedings of the National Academy of Sciences 1994;91:2805–9. https://doi.org/10.1073/pnas.91.7.2805.
[52] Janzen DH. Herbivores and the number of tree species in tropical forests. The American Naturalist 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. Dynamics of Populations 1971;298.
[54] Bagchi R, Gallery RE, Gripenberg S, Gurr SJ, Narayan L, Addis CE, et al. Pathogens and insect herbivores drive rainforest plant diversity and composition. Nature 2014;506:85–8. https://doi.org/10.1038/nature12911.
[55] Raven PH, Gereau RE, Phillipson PB, Chatelain C, Jenkins CN, Ulloa Ulloa C. The distribution of biodiversity richness in the tropics. Sci Adv 2020;6:eabc6228. https://doi.org/10/ghmhwh.
[56] Morley RJ, Morley HP. The prelude to the Holocene: tropical Asia during the Pleistocene. Holocene Climate Change and Environment 2022:1–32.
[57] Morley RJ. The establishment of Palaeotropical rainforests from Africa to Oceania in relation to plate tectonics and zonal tropical climates. Geological Society, London, Special Publications 2025;549:SP549-2023–73. https://doi.org/10.1144/SP549-2023-73.
[58] Morley RJ. The establishment of Palaeotropical rainforests from Africa to Oceania in relation to plate tectonics and zonal tropical climates. Geological Society, London, Special Publications 2024;549:SP549-2023–73. https://doi.org/10.1144/SP549-2023-73.
[59] Hagen O, Skeels A, Onstein RE, Jetz W, Pellissier L. Earth history events shaped the evolution of uneven biodiversity across tropical moist forests. Proc Natl Acad Sci USA 2021;118:e2026347118. https://doi.org/10.1073/pnas.2026347118.
[60] Currano ED, Jacobs BF, Pan AD. Is Africa Really an “Odd Man Out”? Evidence for Diversity Decline across the Oligocene-Miocene Boundary. International Journal of Plant Sciences 2021;182:551–63. https://doi.org/10.1086/714308.
[61] Antonelli A, Nylander JAA, Persson C, Sanmartín I. Tracing the impact of the Andean uplift on Neotropical plant evolution. Proceedings of the National Academy of Sciences 2009;106:9749–54. https://doi.org/10.1073/pnas.0811421106.
[62] Hoorn C, Wesselingh FP, ter Steege H, Bermudez MA, Mora A, Sevink J, et al. Amazonia Through Time: Andean Uplift, Climate Change, Landscape Evolution, and Biodiversity. Science 2010;330:927–31. https://doi.org/10.1126/science.1194585.
[63] Hagen O, Skeels A, Onstein RE, Jetz W, Pellissier L. Earth history events shaped the evolution of uneven biodiversity across tropical moist forests. Proceedings of the National Academy of Sciences 2021;118:e2026347118. https://doi.org/10/gm3gnk.
[64] 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. https://doi.org/10.1007/BF02857848.
[65] Polato NR, Gill BA, Shah AA, Gray MM, Casner KL, Barthelet A, et al. Narrow thermal tolerance and low dispersal drive higher speciation in tropical mountains. Proceedings of the National Academy of Sciences 2018;115:12471–6. https://doi.org/10/ghtfb2.
[66] Levi T, Barfield M, Barrantes S, Sullivan C, Holt RD, Terborgh J. Tropical forests can

maintain hyperdiversity because of enemies. Proceedings of the National Academy of Sciences 2019;116:581–6. https://doi.org/10.1073/pnas.1813211116.
[67] Freeman BG, Strimas-Mackey M, Miller ET. Interspecific competition limits bird species' ranges in tropical mountains. Science 2022;377:416–20. https://doi.org/10.1126/science.abl7242.
[68] Cooper DLM, Lewis SL, Sullivan MJP, Prado PI, Ter Steege H, Barbier N, et al. Consistent patterns of common species across tropical tree communities. Nature 2024;625:728–34. https://doi.org/10.1038/s41586-023-06820-z.
[69] Rubel F, Kottek M. Observed and projected climate shifts 1901-2100 depicted by world maps of the Köppen-Geiger climate classification. Meteorologische Zeitschrift 2010;19:135.
[70] Jost L. The Relation between Evenness and Diversity. Diversity 2010;2:207–32. https://doi.org/10.3390/d2020207.
[71] Hsieh T, Ma Kh, Chao A. iNEXT: an R package for rarefaction and extrapolation of species diversity (H ill numbers). Methods in Ecology and Evolution 2016;7:1451–6.
[72] McGill BJ, Etienne RS, Gray JS, Alonso D, Anderson MJ, Benecha HK, et al. Species abundance distributions: moving beyond single prediction theories to integration within an ecological framework. Ecology Letters 2007;10:995–1015. https://doi.org/10.1111/j.1461-0248.2007.01094.x.
[73] Kursa MB, Rudnicki WR. Boruta: Wrapper Algorithm for All Relevant Feature Selection 2009:8.0.0. https://doi.org/10.32614/CRAN.package.Boruta.
[74] Breiman L. Random Forests. Machine Learning 2001;45:5–32. https://doi.org/10.1023/A:1010933404324.
[75] Aiello-Lammens ME, Boria RA, Radosavljevic A, Vilela B, Anderson RP. 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.
[76] Lê S, Josse J, Husson F. **FactoMineR** : An *R* Package for Multivariate Analysis. J Stat Soft 2008;25. https://doi.org/10.18637/jss.v025.i01.
[77] Kassambara A, Mundt F. Factoextra: extract and visualize the results of multivariate data analyses. CRAN: Contributed Packages 2016.
[78] Benito B. SpatialRF: easy spatial regression with random forest. R Package Version 2021;1.
[79] Elmotawakkil A, Sadiki A, Enneya N. Predicting groundwater level based on remote sensing and machine learning: a case study in the Rabat-Kénitra region. Journal of Hydroinformatics 2024;26:2639–67.
[80] Saltelli A, Ratto M, Andres T, Campolongo F, Cariboni J, Gatelli D, et al. Global sensitivity analysis: the primer. John Wiley & Sons; 2008.

Figures

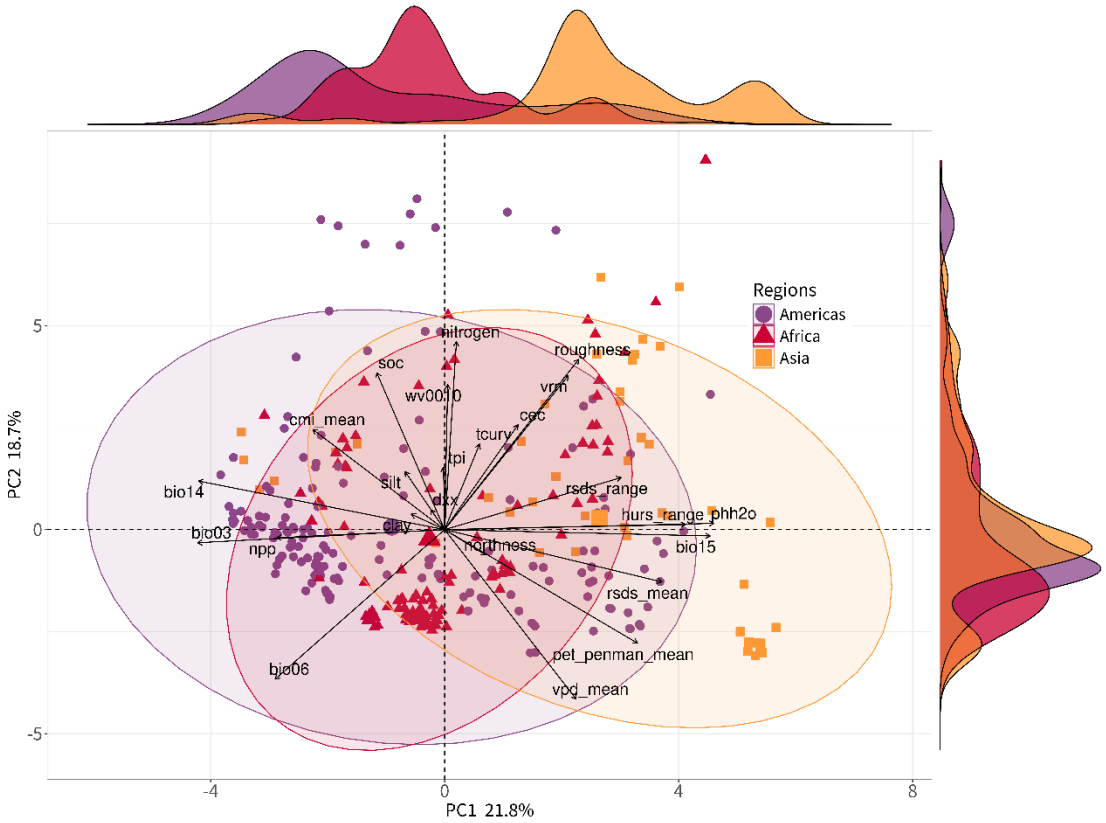

Fig. 1. Principal component analysis (PCA) of 24 environmental variables for 1-ha tropical tree plots in the Americas, Africa and Asia. Marginal density plots above and to the right of the biplot show the distribution of samples from each region along the first (PC1) and second (PC2) principal components,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

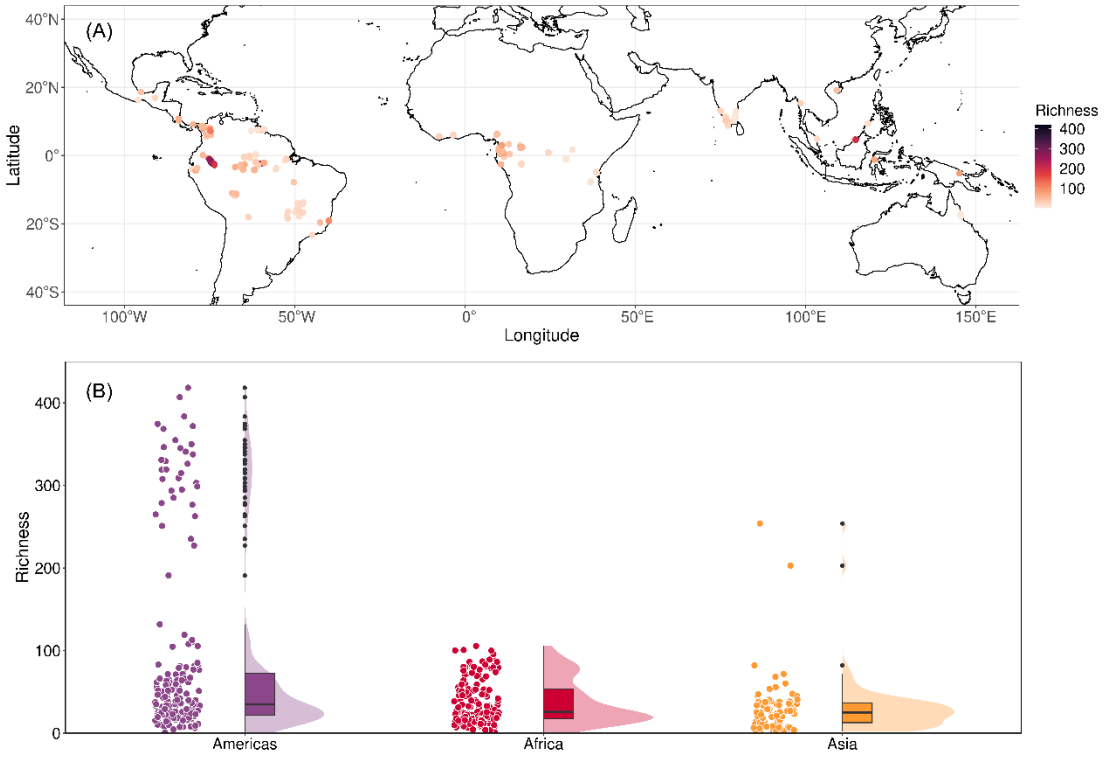

Fig. 2. Tree species richness in 1-ha plots across the global tropics. (**A**) Distribution of 429 1-ha plots located in undisturbed old-growth forests used in this study. Different colored points show tree species richness estimated from rarefaction based on sample coverage. (**B**) Density distribution 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 quartiles and black points represent 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 Supplementary Figs. S3-S4.

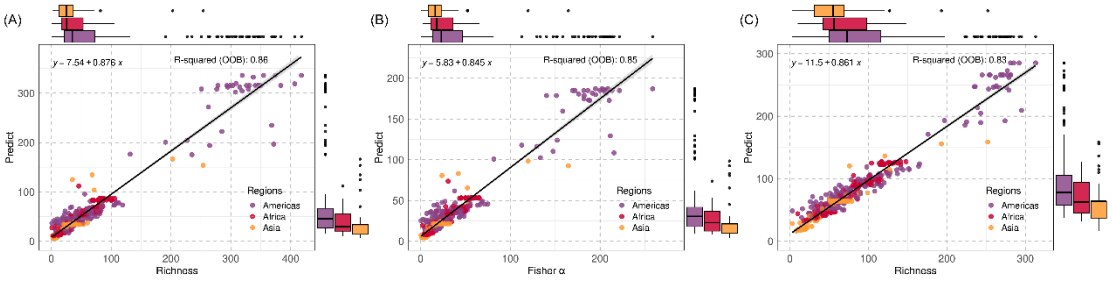

Fig. 3. Predicted tree species richness according to non-spatial Random Forest based on 24 environmental variables compared to richness estimated from (A) sample-based rarefaction, (B) Fisher's α and (C) observed richness. Tree inventory data were obtained from 429 1-ha plots of old-growth tropical forest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

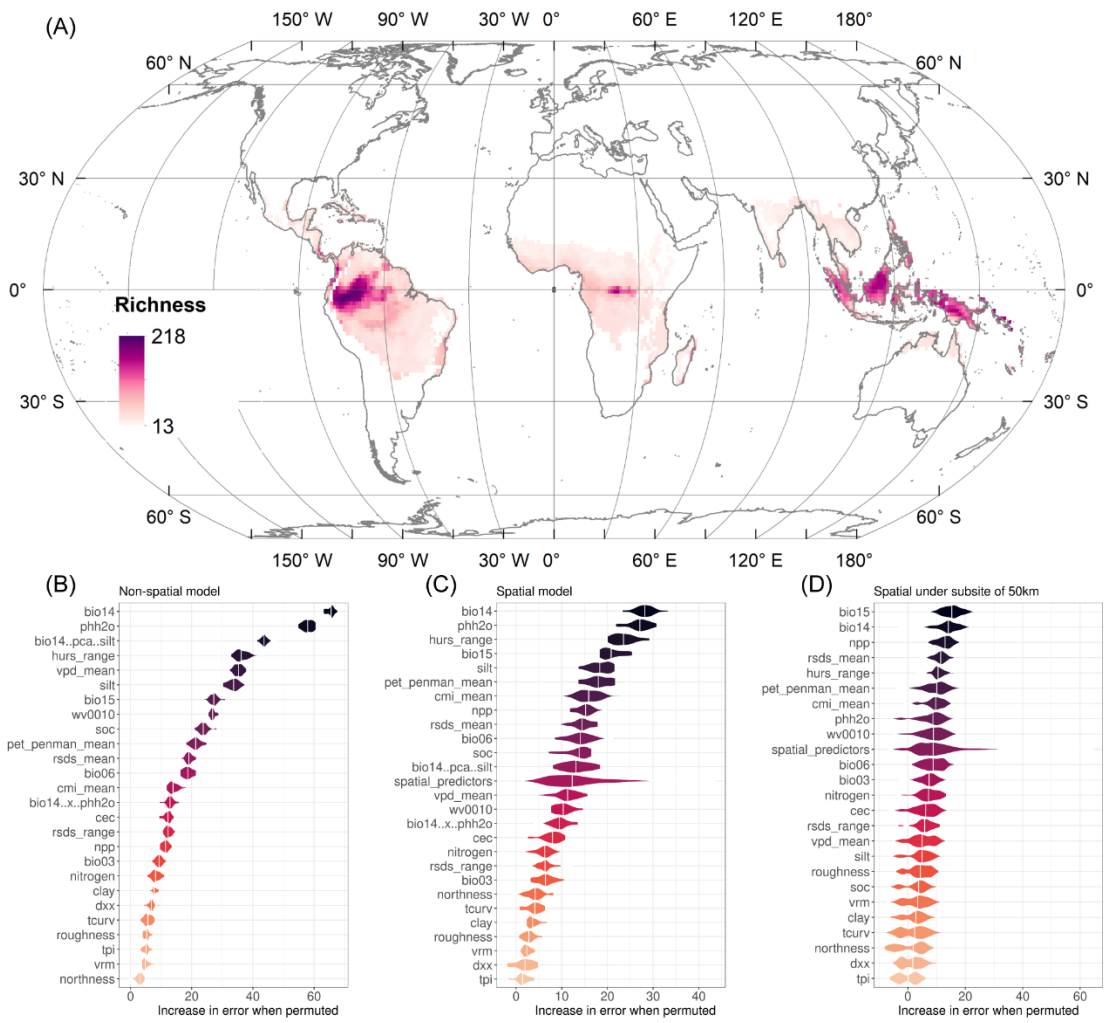

Fig. 4. Predicted tree species richness at the 1-ha scale and importance of environmental variables based on non-spatial and spatial random forest (RF). (**A**) Local tree species richness estimated from rarefaction and predicted using non-spatial RF across the tropics. (**B**) Importance of environmental 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.

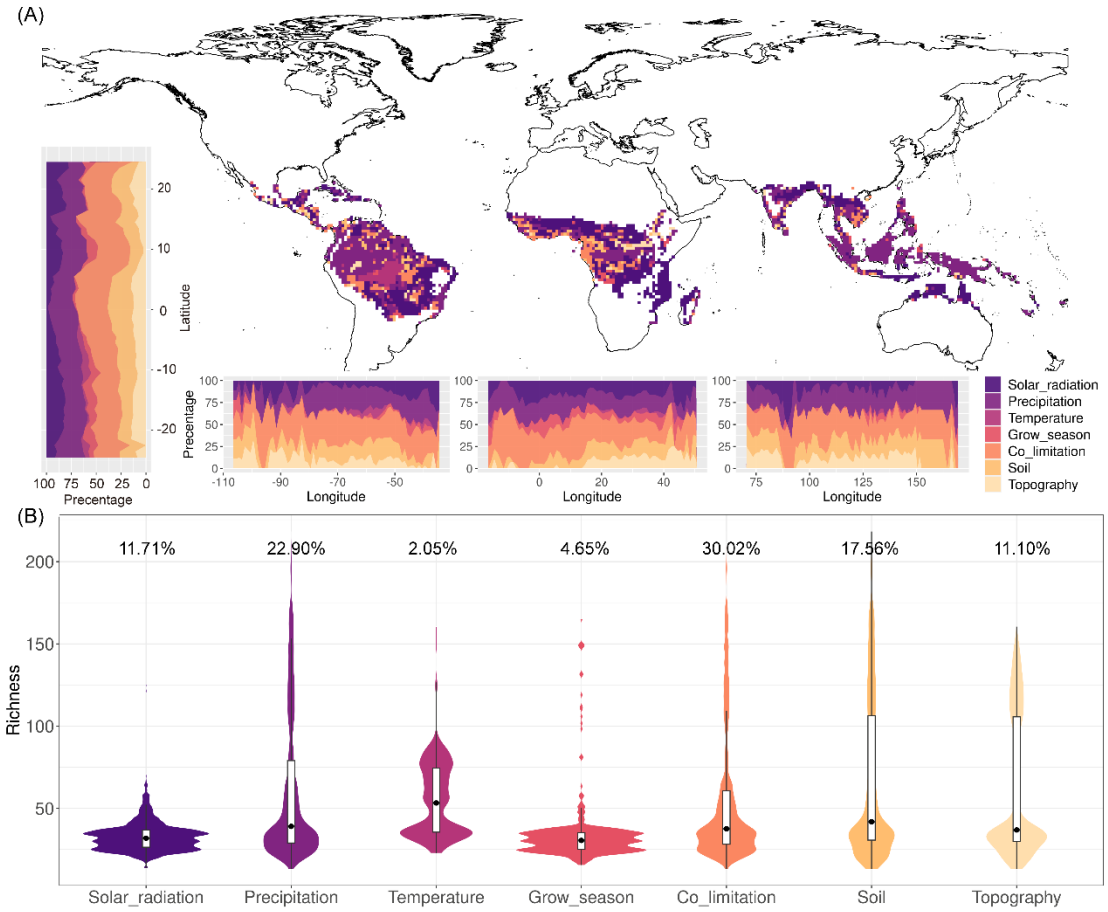

Fig. 5. Geographical distribution of predominance of seven categories of environmental variables. Colimitation refers to areas where no single factor dominates. (**A**) The driving factors distribution pattern in pan-tropical regions, along longitude and latitude. (**B**) The percentage of main categories driving factors in tropical region. A detailed variables for different drivers were explained in Table S3.