
Article:

Liang, J, Gamarra, JGP, Picard, N et al. (2022) Co-limitation towards lower latitudes shapes global forest diversity gradients. Nature Ecology and Evolution, 6 (10). pp. 1423-1437. ISSN: 2397-334X

https://doi.org/10.1038/s41559-022-01831-x

Co-limitation toward lower latitudes shapes global forest diversity gradients

Jingjing Liang[1]*, Javier G.P. Gamarra[2], Nicolas Picard[3], Mo Zhou[1], Bryan Pijanowski[4], Douglass F. Jacobs[4], Peter B. Reich[5,6,7], Thomas W. Crowther[8], Gert-Jan Nabuurs[9,10], Sergio de-Miguel[11,12], Jingyun Fang[13], Christopher W. Woodall[14], Jens-Christian Svenning[15,16], Tommaso Jucker[17], Jean-Francois Bastin[18], Susan K. Wiser[19], Ferry Slik[20], Bruno Hérault[21,22], Giorgio

---

[1] Forest Advanced Computing and Artificial Intelligence Laboratory (FACAI), Department of Forestry and Natural Resources, Purdue University, West Lafayette, IN 47907, USA

[2] Forestry Division, Food and Agriculture Organization of the United Nations, Rome 00153, Italy

[3] GIP ECOFOR, Paris 75116, France

[4] Department of Forestry and Natural Resources, Purdue University, West Lafayette, IN 47907, USA

[5] Institute for Global Change Biology, School for Environment and Sustainability, University of Michigan, Ann Arbor, MI 48109, USA

[6] Department of Forest Resources, University of Minnesota, St. Paul, MN 55108, USA

[7] Hawkesbury Institute for the Environment, Western Sydney University, Penrith, NSW 2753, Australia

[8] Crowther Lab, Department of Environmental Systems Science, Institute of Integrative Biology, ETH Zürich, Zürich 8092, Switzerland

[9] Wageningen Environmental Research, Wageningen University and Research, NL6700AA, the Netherlands

[10] Forest Ecology and Forest Management Group, Wageningen University and Research, NL6700AA, the Netherlands

[11] Department of Crop and Forest Sciences, University of Lleida, Lleida E25198, Spain

[12] Joint Research Unit CTFC - Agrotecnio - CERCA, Solsona E25280, Spain

[13] Institute of Ecology and Key Laboratory for Earth Surface Processes of the Ministry of Education, College of Urban and Environmental Sciences, Peking University, Beijing 100871, China

[14] Northern Research Station, USDA Forest Service, Durham, NH 03824, USA

[15] Center for Biodiversity Dynamics in a Changing World (BIOCHANGE), Department of Biology, Aarhus University, DK-8000 Aarhus C, Denmark

[16] Section for Ecoinformatics and Biodiversity, Department of Biology, Aarhus University, DK-8000 Aarhus C, Denmark

[17] School of Biological Sciences, University of Bristol, Bristol BS8 1TQ, UK

[18] TERRA Teaching and Research Centre, Gembloux Agro Bio-Tech, University of Liege, 5030 Gembloux, Belgium

[19] Manaaki Whenua Landcare Research PO Box 40, Lincoln 7640, New Zealand

[20] Environmental and Life Sciences, Faculty of Science, Universiti Brunei Darussalam, Gadong BE1410, Brunei Darussalam

[21] Centre de Coopération Internationale en Recherche Agronomique pour le Développement, Montpellier 34398, France

[22] INP-HB (Institut National Polytechnique Félix Houphouet-Boigny), Univ Montpellier, BP 1093, Yamoussoukro, Ivory Coast

Alberti[23,24,25], Gunnar Keppel[26], Geerten M. Hengeveld[27,28], Pierre L. Ibisch[29], Carlos A. Silva[30], Hans H. ter Steege[31], Pablo L. Peri[32], David A. Coomes[33], Eric B. Searle[34], Klaus von Gadow[35,36,37], Bogdan Jaroszewicz[38], Akane O. Abbasi[1], Meinrad Abegg[39], Yves C. Adou Yao[40], Jesús Aguirre-Gutiérrez[41,42], Angelica M. Almeyda Zambrano[43], Jan Altman[44,45], Esteban Alvarez-Dávila[46], Juan Gabriel Álvarez-González[47], Luciana F. Alves[48], Bienvenu H.K. Amani[49], Christian A. Amani[50], Christian Ammer[51], Bhely Angoboy Ilondea[52], Clara Antón-Fernández[53], Valerio Avitabile[54], Gerardo A. Aymard[55], Akomian F. Azihou[56], Johan A Baard[57],

---

[23] Department of Agricultural, Food, Environmental and Animal Sciences, University of Udine, Udine 33100, Italy
[24] Faculty of Science and Technology, Free University of Bolzano, Bolzano 39100, Italy
[25] Institute of Bioeconomy, CNR, Sesto 50019, Italy
[26] Natural and Built Environments Research Centre, School of Natural and Built Environments University of South Australia, Adelaide, SA 5001, Australia
[27] Biometris, Wageningen University and Research, Wageningen 6708, the Netherlands
[28] Wageningen University & Research, Forest and Nature Conservation Policy Group, Wageningen 6708, the Netherlands
[29] Centre for Econics and Ecosystem Management, Eberswalde University for Sustainable Development, Eberswalde 16225, Germany
[30] School of Forest, Fisheries, and Geomatics Sciences, Institute of Food & Agricultural Sciences, University of Florida, Gainesville, FL 32611, USA
[31] Naturalis Biodiversity Center, Leiden 2333, the Netherlands
[32] Instituto Nacional de Tecnología Agropecuaria (INTA), Santa Cruz Z9400, Argentina
[33] Department of Plant Sciences, University of Cambridge, Cambridge CB23EA, UK
[34] Faculty of Natural Resources Management, Lakehead University, Thunder Bay, ON P7B 5E1, Canada
[35] University of Göttingen, Göttingen 37073, Germany
[36] Beijing Forestry University, Beijing 100107, China
[37] University of Stellenbosch, Stellenbosch, South Africa
[38] Białowieża Geobotanical Station, Faculty of Biology, University of Warsaw, Białowieża PL-17-230, Poland
[39] Swiss National Forest Inventory /Swiss Federal Institute for Forest, Snow and Landscape Research WSL, Birmensdorf CH-8903, Switzerland
[40] UFR Biosciences, University Félix Houphouët-Boigny, Abidjan, Ivory Coast
[41] Environmental Change Institute, School of Geography and the Environment, University of Oxford, Oxford OX1 3QY, UK
[42] Biodiversity Dynamics, Naturalis Biodiversity Center, Leiden 2333, the Netherlands
[43] Center for Latin American Studies, University of Florida, Gainesville, FL 32611, USA
[44] Institute of Botany, Academy of Sciences of the Czech Republic, Trebon 379 01, Czech Republic
[45] Faculty of Forestry and Wood Sciences, Czech University of Life Sciences in Prague, Praha-Suchdol 16521, Czech Republic
[46] Escuela ECAPMA, National Open University and Distance (Colombia) | UNAD, Bogotá, Colombia
[47] Departamento de Ingeniería Agroforestal, Universidad de Santiago de Compostela, Lugo 27002, Spain
[48] Center for Tropical Research, Institute of the Environment and Sustainability, University of California, Los Angeles, CA 90095, USA
[49] Université Jean Lorougnon Guédé, Daloa, Côte d'Ivoire
[50] Université Officielle de Bukavu, Bukavu, Democratic Republic of Congo
[51] Silviculture and Forest Ecology of the Temperate Zones, University of Göttingen, Goettingen 37077, Germany
[52] Institut National pour l'Etude et la Recherche Agronomiques, Kinshasa, Democratic Republic of Congo
[53] Norwegian Institute of Bioeconomy Research (NIBIO), Division of Forestry and Forest Resources, Ås NO-1431, Norway
[54] European Commission, Joint Research Centre, Ispra 21027, Italy
[55] Compensation International Progress S.A., P.O. Box 260161, Bogotá, D. C., Colombia
[56] Laboratory of Applied Ecology, University of Abomey-Calavi, 01 BP 526 Cotonou, Benin
[57] Scientific Services, South African National Parks, Knysna 6570, South Africa

Timothy R Baker[58], Radomir Balazy[59], Meredith L. Bastian[60,61], Rodrigue Batumike[62], Marijn Bauters[63,64], Hans Beeckman[65], Nithanel Mikael Hendrik Benu[66], Robert Bitariho[67], Pascal Boeckx[64], Jan Bogaert[68], Frans Bongers[10], Olivier Bouriaud[69], Pedro H.S. Brancalion[70], Susanne Brandl[71], Francis Q. Brearley[72], Jaime Briseno-Reyes[73], Eben N. Broadbent[30], Helge Bruelheide[74,75], Erwin Bulte[76], Ann Christine Catlin[77], Roberto Cazzolla Gatti[78], Ricardo G. César[70], Han Y.H. Chen[34], Chelsea Chisholm[79], Emil Cienciala[80,81], Gabriel D. Colletta[82], José Javier Corral-Rivas[73], Anibal Cuchietti[83], Aida Cuni-Sanchez[84,85], Javid A. Dar[86,87,88], Selvadurai

---

[58] School of Geography, University of Leeds, Leeds LS2 9JT, UK
[59] Department of Geomatics, Forest Research Institute, Sekocin Stary, 05-090 Raszyn, Poland
[60] Proceedings of the National Academy of Sciences, Washington, DC 20001, USA
[61] Department of Evolutionary Anthropology, Duke University, Durham, NC 27708, USA
[62] Department of Environment, Universtité du Cinquantenaire de Lwiro, Bukavu, Democratic Republic of Congo
[63] Department of Environment, Ghent University, Gent 9000, Belgium
[64] Department of Green Chemistry and Technology, Ghent University, Gent 9000, Belgium
[65] Service of Wood Biology, Royal Museum for Central Africa, Tervuren 3080, Belgium
[66] Balai Penelitian dan Pengembangan Lingkungan Hidup dan Kehutanan (BP2LHK), Manokwari 98314, Indonesia
[67] Institute of Tropical Forest Conservation, Mbarara University of Science and Technology, Mbarara, Uganda

[68] Université de Liège, Gembloux Agro-Bio Tech, Gembloux, Belgium
[69] Integrated Center for Research, Development and Innovation in Advanced Materials, Nanotechnologies, and Distributed Systems for Fabrication and Control (MANSiD), University Stefan cel Mare of Suceava, Suceava 720229, Romania
[70] Department of Forestry Sciences, "Luiz de Queiroz" College of Agriculture, University of São Paulo, Piracicaba 13400-970, Brazil
[71] Bavarian State Institute of Forestry, Freising 85354, Germany
[72] Department of Natural Sciences, Manchester Metropolitan University, Manchester M1 5GD, UK
[73] Facultad de Ciencias Forestales, Universidad Juárez del Estado de Durango, Durango 34120, Mexico
[74] Institute of Biology and Botanical Garden, Martin Luther University Halle-Wittenberg, Halle (Saale) 06108, Germany
[75] German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Leipzig 04103, Germany
[76] Development Economics Group, Wageningen University, Wageningen 6706KN, the Netherlands
[77] Rosen Center for Advanced Computing (RCAC), Purdue University, West Lafayette, IN 47907, USA
[78] Department of Biological, Geological and Environmental Sciences (BiGeA), University of Bologna, Bologna 40, Italy
[79] Department of Integrative Biology, ETH Zürich, Zürich 8092, Switzerland
[80] Institute of Forest Ecosystem Research, Jilove u Prahy CZ 254 01, Czech Republic
[81] Global Change Research Institute of the CAS, Brno CZ 603 00, Czech Republic
[82] Programa de Pós-graduação em Biologia Vegetal, Instituto de Biologia, Universidade Estadual de Campinas, Campinas CEP 13083-862, Brazil
[83] Dirección Nacional de Bosques (DNB), Ministerio de Ambiente y Desarrollo Sostenible (MAyDS), Ciudad Autónoma de Buenos Aires C1004AAI, Argentina
[84] Department of International Environment and Development Studies (Noragric), Faculty of Landscape and Society, Norwegian University of Life Sciences (NMBU), Ås N-1432, Norway
[85] Department of Environment and Geography, University of York, York YO10 5NG, UK
[86] Department of Environmental Science, School of Engineering and Sciences, SRM University-AP, Guntur 522240, India
[87] Department of Botany, Dr. Harisingh Gour Vishwavidyalaya (A Central University), Madhya Pradesh-470003, India
[88] Department of Ecology and Environmental Sciences, Pondicherry University, Puducherry-605014, India

Dayanandan[89], Thales de Haulleville[65,68], Mathieu Decuyper[10], Sylvain Delabye[90,91], Géraldine Derroire[92], Ben DeVries[93], John Diisi[94], Tran Van Do[95], Jiri Dolezal[44,96], Aurélie Dourdain[92], Graham P. Durrheim[97], Nestor Laurier Engone Obiang[98], Corneille E.N. Ewango[99], Teresa J. Eyre[100], Tom M. Fayle[101,91], Lethicia Flavine N. Feunang[102], Leena Finér[103], Markus Fischer[104], Jonas Fridman[105], Lorenzo Frizzera[106], André L. de Gasper[107], Damiano Gianelle[108], Henry B. Glick[109], Maria Socorro Gonzalez-Elizondo[110], Lev Gorenstein[77], Richard Habonayo[111], Olivier J. Hardy[112], David J. Harris[113], Andrew Hector[114], Andreas Hemp[115], Martin Herold[116], Annika

---

[89] Centre for Structural and Functional Genomics & Quebec Centre for Biodiversity Science, Biology Department, Concordia University, Montreal, QC H4B 1R6, Canada
[90] Department of Ecology, Faculty of Science, Charles University, Ceske Budejovice 12844, Czech Republic
[91] Biology Centre of the Czech Academy of Sciences, Institute of Entomology, Ceske Budejovice 37005, Czech Republic
[92] Cirad, UMR EcoFoG (AgroParistech, CNRS, Inrae, Université des Antilles, Université de la Guyane), Campus Agronomique, Kourou 97387, French Guiana
[93] Department of Geography, Environment and Geomatics, University of Guelph, Guelph, ON N1G 2W1, Canada
[94] National Forest Authority, Kampala, Uganda
[95] Department of Silviculture Foundation, Silviculture Research Institute, Vietnamese Academy of Forest Sciences, Hanoi, Vietnam
[96] Department of Botany, Faculty of Science, University of South Bohemia, Ceske Budejovice 370 05, Czech Republic
[97] Scientific Services, South African National Parks, Knysna 6570, South Africa
[98] IPHAMETRA, IRET, CENAREST, Libreville, Gabon
[99] Faculté de Gestion de Ressources Naturelles Renouvelables, Université de Kisangani, Kisangani R408, Democratic Republic of Congo
[100] Queensland Herbarium, Department of Environment and Science, Toowong, QLD 4066, Australia
[101] School of Biological and Behavioural Sciences, Queen Mary University of London, London E1 4NS, UK
[102] Department of Plant Biology, Faculty of Science, University of Yaoundé I, Yaoundé, Cameroon
[103] Natural Resources Institute Finland, Joensuu FI-80100, Finland
[104] Institute of Plant Sciences, University of Bern, Bern CH-3013, Switzerland
[105] Department of Forest Resource Management, Swedish University of Agricultural Sciences, Umea SE-90183, Sweden
[106] Department of Sustainable Agro-Ecosystems and Bioresources, Research and Innovation Center, Trento 38010, Italy
[107] Herbário Dr. Roberto Miguel Klein, Universidade Regional de Blumenau, Rua Antônio da Veiga, Blumenau, SC 89030-903, Brazil
[108] Department of Sustainable Agro-ecosystems and Bioresources, Research and Innovation Centre, Fondazione Edmund Mach, Via E. Mach 1, 38010 San Michele all'Adige (TN), Italy
[109] Glick Designs, LLC., Hadley, MA 01035, USA
[110] CIIDIR Durango, Instituto Politécnico Nacional, Durango 34120, Mexico
[111] Département des Sciences et Technologies de l'Environnement, Université du Burundi, Bujumbura BP 2940, Burundi
[112] Faculté des Sciences, Evolutionary Biology and Ecology Unit, Université Libre de Bruxelles, Brussels 1050, Belgium
[113] Royal Botanic Garden Edinburgh, Edinburgh EH3 5LR, UK
[114] Department of Plant Sciences, University of Oxford, Oxford OX1 3RB, UK
[115] Department of Plant Systematics, Bayreuth University, Bayreuth 95440, Germany
[116] Helmholtz GFZ German Research Centre for Geosciences, Section 1.4 Remote Sensing and Geoinformatics, Potsdam 14473, Germany

Hillers[117,118], Wannes Hubau[119,120], Thomas Ibanez[121], Nobuo Imai[122], Gerard Imani[123], Andrzej M. Jagodzinski[124,125], Stepan Janecek[90], Vivian Kvist Johannsen[126], Carlos A. Joly[127], Blaise Jumbam[128,129], Banoho L.P.R. Kabelong[102], Goytom Abraha Kahsay[130], Viktor Karminov[131], Kuswata Kartawinata[132], Justin N. Kassi[133], Elizabeth Kearsley[134], Deborah K. Kennard[135], Sebastian Kepfer-Rojas[126], Mohammed Latif Khan[136], John N. Kigomo[137], Hyun Seok Kim[138,139,140,141], Carine Klauberg[30], Yannick Klomberg[90], Henn Korjus[142], Subashree Kothandaraman[88,87], Florian Kraxner[143], Amit Kumar[144], Relawan Kuswandi[66], Mait Lang[142,145], Michael J. Lawes[146], Rodrigo V. Leite[147], Geoffrey Lentner[77], Simon L. Lewis[58,148], Moses B.

---

[117] Wild Chimpanzee Foundation, Liberia Representation, FDA Compound, Mt. Barclay, Liberia
[118] Centre for Conservation Science, The Royal Society for the Protection of Birds, Sandy SG19 2DL, UK
[119] Department of Environment, Laboratory for Wood Technology (UGent-Woodlab), Ghent University, Gent 9000, Belgium
[120] Service of Wood Biology, Royal Museum for Central Africa, Tervuren 3080, Belgium
[121] AMAP, Univ Montpellier, CIRAD, CNRS, INRA, IRD, Montpellier 34000, France
[122] Department of Forest Science, Tokyo University of Agriculture, Tokyo 1568502, Japan
[123] Biology department, Université Officielle de Bukavu, Bukavu, Democratic Republic of Congo
[124] Institute of Dendrology, Polish Academy of Sciences, Kórnik PL-62-035, Poland
[125] Poznan University of Life Sciences, Faculty of Forestry and Wood Technology, Department of Game Management and Forest Protection, Poznan PL-60-625, Poland
[126] Department of Geosciences and Natural Resource Management, University of Copenhagen, Copenhagen 1958, Denmark
[127] Plant Biology Department, Biology Institute, University of Campinas (UNICAMP), Campinas, SP 13083-970, Brazil
[128] Department of Botany and Plant Pathology, Purdue University, West Lafayette, IN 47907, USA
[129] Institute of Agricultural Research for Development (IRAD), Nkolbisson, Ministry of Scientific Research and Innovation, Yaounde, Cameroon
[130] Department of Food and Resource Economics, University of Copenhagen, Copenhagen DK-1958, Denmark
[131] Forestry Faculty, Bauman Moscow State Technical University, Mytischi 141005, Russia
[132] Integrative Research Center, The Field Museum, Chicago, IL 60605, USA
[133] Labo botanique, Université Félix Houphouët-Boigny, Abidjan, Côte d'Ivoire
[134] Computational and Applied Vegetation Ecology lab, Ghent University, Gent 9000, Belgium
[135] Department of Physical and Environmental Sciences, Colorado Mesa University, Grand Junction, CO 81501, USA
[136] Department of Botany, Dr. Harisingh Gour Vishwavidalaya (A Central University), Sagar, MP 470003, India
[137] Kenya Forestry Research Institute, Department of Forest Resource Assessment, Nairobi, Kenya
[138] Department of Forest Sciences, Seoul National University, Seoul 08826, Republic of Korea
[139] Interdisciplinary Program in Agricultural and Forest Meteorology, Seoul National University, Seoul 08826, Republic of Korea
[140] National Center for Agro Meteorology, Seoul 08826, Republic of Korea
[141] Research Institute for Agriculture and Life Sciences, Seoul National University, Seoul 08826, Republic of Korea
[142] Institute of Forestry and Engineering, Estonian University of Life Sciences, Tartu 51006, Estonia
[143] International Institute for Applied Systems Analysis, Laxenburg A-2361, Austria
[144] Department of Geoinformatics, Central University of Jharkhand, Ranchi, Jharkhand 835205, India
[145] Tartu Observatory, University of Tartu, Tõravere 61602, Estonia
[146] School of Life Sciences, University of KwaZulu-Natal, Pietermaritzburg 3209, South Africa
[147] Department of Forest Engineering, Federal University of Viçosa (UFV), Viçosa, MG 36570-900, Brazil
[148] Department of Geography, University College London, London WC1E 6BT, UK

Libalah[149,102] , Janvier Lisingo[150], Pablito Marcelo López-Serrano[151], Huicui Lu[152], Natalia V. Lukina[153], Anne Mette Lykke[154], Vincent Maicher[90,91,155], Brian S. Maitner[156], Eric Marcon[92], Andrew R. Marshall[157,158,159], Emanuel H. Martin[160], Olga Martynenko[131], Faustin M. Mbayu[99], Musingo T.E. Mbuvi[161], Jorge A. Meave[162], Cory Merow[163], Stanislaw Miscicki[164], Vanessa S. Moreno[70], Albert Morera[12,12], Sharif A. Mukul[165], Jörg C. Müller[166,167], Agustinus Murdjoko[168], Maria Guadalupe Nava-Miranda[151], Litonga Elias Ndive[169], Victor J. Neldner[100], Radovan V. Nevenic[170], Louis N. Nforbelie[102], Michael L. Ngoh[171,172], Anny E. N'Guessan[40], Michael R. Ngugi[100], Alain S.K. Ngute[165,173], Emile Narcisse N. Njila[102], Melanie C. Nyako[102], Thomas O. Ochuodho[174], Jacek Oleksyn[124], Alain Paquette[175], Elena I. Parfenova[176], Minjee Park[4], Marc Parren[10], Narayanaswamy Parthasarathy[88], Sebastian Pfautsch[177], Oliver L. Phillips[58], Maria T.F.

---

[149] Plant Systematics and Ecology Laboratory (LaBosystE), Higher Teacher's Training College, University of Yaoundé I, Yaoundé, Cameroon

[150] Faculté des Sciences, Laboratoire d'écologie et aménagement forestier, Université de Kisangani, Kisangani, Democratic Republic of Congo

[151] Instituto de Silvicultura e Industria de la Madera, Universidad Juarez del Estado de Durango, Durango 34120, Mexico

[152] Faculty of Forestry, Qingdao Agricultural University, Qingdao 266000, China

[153] Center for Forest Ecology and Productivity RAS (CEPF RAS), Moscow, Russian Federation

[154] Department of Ecoscience, Aarhus University, Silkeborg 8600, Denmark

[155] Nicholas School of the Environment, Duke University, Durham, NC 27710, USA

[156] Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, AZ 85721, USA

[157] University of the Sunshine Coast, Sippy Downs, QLD 4556, Australia

[158] University of York, York YO10 5NG, UK

[159] Flamingo Land Ltd., North Yorkshire YO10 6UX, UK

[160] Department of Wildlife Management, College of African Wildlife Management, Moshi, Tanzania

[161] Kenya Forestry Research Institute, Headquarters, Nairobi, Kenya

[162] Departamento de Ecología y Recursos Naturales, Facultad de Ciencias, Universidad Nacional Autónoma de México, Ciudad de México 04510, Mexico

[163] Ecology and Evolutionary Biology, University of Connecticut, Storrs, CT 06268, USA

[164] Department of Forest Management and Forest Economics, Warsaw University of Life Sciences, Warsaw 02-776, Poland

[165] Tropical Forests and People Research Centre, University of the Sunshine Coast, Maroochydore DC, QLD 4558, Australia

[166] Fieldstation Fabrikschleichach, Julius-Maximilians University Würzburg, Würzburg 97070, Germany

[167] Bavarian Forest Nationalpark, Grafenau 94481, Germany

[168] Fakultas Kehutanan, Universitas Papua, Jalan Gunung Salju Amban 98314, Manokwari Papua Barat, Indonesia

[169] Limbe Botanic Garden, Limbe, Cameroon

[170] Institute of Forestry, Belgrade 11000, Serbia

[171] Tropical Plant Exploration Group (TroPEG), Buea, Cameroon

[172] Department of Ecology and Evolutionary Biology, University of Tennessee, Knoxville, TN 37996, USA

[173] Applied Biology and Ecology Research Unit, University of Dschang, Dschang, Cameroon

[174] Department of Forestry and Natural Resources, University of Kentucky, Lexington, KY 40546, USA

[175] UQAM, Centre for forest research, Montreal, QC H3C 3P8, Canada

[176] V.N. Sukachev Forest Institute of FRC KSC SB RAS, Krasnoyarsk 660036, Russia

[177] Urban Management and Planning, School of Social Sciences, Western Sydney University, Penrith, NSW 2751, Australia

Piedade[178], Daniel Piotto[179], Martina Pollastrini[180], Lourens Poorter[10], John R. Poulsen[155], Axel Dalberg Poulsen[113], Hans Pretzsch[181], Mirco Rodeghiero[106,182], Samir G. Rolim[179], Francesco Rovero[183,184], Ervan Rutishauser[185], Khosro Sagheb-Talebi[186], Purabi Saikia[187], Moses Nsanyi Sainge[171,188], Christian Salas-Eljatib[189,190,191], Antonello Salis[2], Peter Schall[51], Dmitry Schepaschenko[143,192,], Michael Scherer-Lorenzen[193], Bernhard Schmid[194], Jochen Schöngart[178], Vladimír Šebeň[195], Giacomo Sellan[196,72], Federico Selvi[180], Josep M. Serra-Diaz[197], Douglas Sheil[10,198], Anatoly Z. Shvidenko[143], Plinio Sist[199], Alexandre F. Souza[200], Krzysztof J. Station Stereńczak[59], Martin J.P. Sullivan[72], Somaiah Sundarapandian[88], Miroslav Svoboda[45], Mike D. Swaine[201], Natalia Targhetta[178], Nadja Tchebakova[176], Liam A. Trethowan[202], Robert Tropek[90,91], John Tshibamba Mukendi[203], Peter Mbanda Umunay[204], Vladimir A. Usoltsev[205], Gaia Vaglio Laurin[206], Riccardo Valentini[206], Fernando Valladares[207], Fons van der Plas[208], Daniel José Vega-

---

[178] Instituto Nacional de Pesquisas da Amazônia - INPA, Grupo Ecologia. Monitoramento e Uso Sustentável de Áreas Úmidas MAUA, Manaus, Amazonas, Brazil

[179] Centro de Formação em Ciências Agroflorestais, Universidade Federal do Sul da Bahia, Ilhéus, BA 45613-204, Brazil

[180] Dept of Agriculture, Food, Environment and Forestry, University of Firenze, Firenze 50144, Italy

[181] Technical University of Munich, School of Life Sciences Weihenstephan, Chair of Forest Growth and Yield Science, Munich 85354, Germany

[182] Centro Agricoltura, Alimenti, Ambiente, University of Trento, San Michele all'Adige, TN 38010, Italy

[183] Department of Biology, University of Florence, Sesto Fiorentino 50019, Italy

[184] MUSE - Museo delle Scienze, Trento 38122, Italy

[185] Infoflora c/o Botanical Garden of Geneva, Geneva, Switzerland

[186] Agricultural Research, Education and Extension Organization (AREEO), Research Institute of Forests and Rangelands (RIFR), Tehran Q9WX+9WC, Iran

[187] Department of Environmental Sciences, Central University of Jharkhand, Ranchi, Jharkhand 835205, India

[188] Institute of International Education Scholar Rescue Fund (IIE-SRF), One World Trade Center, New York, NY 10007, USA

[189] Centro de Modelación y Monitoreo de Ecosistemas, Facultad de Ciencias, Universidad Mayor, Santiago, Chile

[190] Vicerrectoria de Inv. y Post., Universidad de La Frontera, Temuco, Chile

[191] Depto. de Silvicultura y Conservación de la Naturaleza, Universidad de Chile, Santiago, Chile

[192] V.N. Sukachev Institute of Forest, Siberian Branch of the Russian Academy of Science, Krasnoyarsk 660036, Russia

[193] University of Freiburg, Faculty of Biology, Freiburg D-79104, Germany

[194] Institution with City, Department of Geography, University of Zurich, Zurich CH-8057, Switzerland

[195] National Forest Centre, Zvolen 96001, Slovak Republic

[196] CNRS-UMR LEEISA, Campus Agronomique, Kourou 97310, French Guiana

[197] Universite de Lorraine, AgroParisTech, INRA, Nancy 5400, France

[198] Center for International Forestry Research (CIFOR), Situ Gede, Bogor Barat, Jawa Barat 16115, Indonesia

[199] Cirad, University of Montpellier, Montpellier, France

[200] Universidade Federal do Rio Grande do Norte, Departamento de Ecologia, Natal 59094-010, Brazil

[201] School of Biological Sciences, University of Aberdeen, Aberdeen AB24 3FX, UK

[202] Herbarium Kew, Royal Botanic Gardens Kew, London TW9 3AE, UK

[203] Faculté des Sciences Appliquées, Université de Mbujimayi, Mbujimayi, Democratic Republic of Congo

[204] Yale School of Forestry and Environmental Studies, New Haven, CT 06511, USA

[205] Ural State Forest Engineering University, 620100, Botanical Garden, Ural Branch of the Russian Academy of Sciences, 620144, Yekaterinburg, Russia

[206] DIBAF Department, Tuscia University, Viterbo, Italy

[207] LINCGlobal, MNCN, CSIC, Madrid 28006, Spain

[208] Plant Ecology and Nature Conservation Group, Wageningen University, PO Box 47, 6700, AA Wageningen, the Netherlands

Nieva[73], Hans Verbeeck[134], Helder Viana[209,210], Alexander C. Vibrans[211], Simone A. Vieira[212], Jason Vleminckx[213], Catherine E. Waite[214], Hua-Feng Wang[215], Eric Katembo Wasingya[99], Chemuku Wekesa[216], Bertil Westerlund[105], Florian Wittmann[217], Verginia Wortel[218], Tomasz Zawiła-Niedźwiecki[219], Chunyu Zhang[220], Xiuhai Zhao[220], Jun Zhu[221], Xiao Zhu[77], Zhi-Xin Zhu[215], Irie C. Zo-Bi[222], Cang Hui[223,224*]

*Corresponding authors: Jingjing Liang (albeca.liang@gmail.com), Cang Hui (chui@sun.ac.za)

The latitudinal diversity gradient (LDG) is one of the most recognized global patterns of species richness exhibited across a wide range of taxa. Numerous hypotheses have been proposed in the last two centuries to explain LDG, but rigorous tests of the drivers of LDGs have been limited by a lack of high-quality global species richness data. Here, we produce a high-resolution (0.025°×0.025°) map of local tree species richness using a global forest inventory database with individual tree information and local biophysical characteristics

---

[209] Agricultural High School, ESAV, Polytechnic Institute of Viseu, IPV, Viseu 3500-606, Portugal
[210] Centre for the Research and Technology of Agro-Environmental and Biological Sciences, CITAB, UTAD, Quinta de Prados, Vila Real 5000-801, Portugal
[211] Department of Forest Engineering, Universidade Regional de Blumenau, Blumenau 89030-000, Brazil
[212] Nucleo de Estudos e Pesquisas Ambientais, Universidade Estadual de Campinas, Campinas (UNICAMP), SP 13083-970, Brazil
[213] International Center for Tropical Botany, Department of Biological Sciences, Florida International University, Miami, FL 33199, USA
[214] Forest Research Institute, University of the Sunshine Coast, Sippy Downs, QLD 4556, Australia
[215] Key Laboratory of Tropical Biological Resources of Ministry of Education, School of Life and Pharmaceutical Sciences, Hainan University, Haikou 570228, China
[216] Kenya Forestry Research Institute, Taita Taveta Research Centre, Wundanyi, Kenya
[217] Dep. of Wetland Ecology, Institute for Geography and Geoecology, Karlsruhe Institute for Technology, Rastatt D-76437, Germany
[218] Department of Forest Management, Centre for Agricultural Research in Suriname, Prof. Dr. Ir. J. Ruinardlaan, Paramaribo, Zuid, Suriname
[219] Polish State Forests-Coordination Centre for Environmental Projects, Warsaw 02-362, Poland
[220] Research Center of Forest Management Engineering of State Forestry and Grassland Administration, Beijing Forestry University, Beijing 100083, China
[221] Department of Statistics, University of Wisconsin - Madison, Madison, WI 53706, USA
[222] Institut National Polytechnique Félix Houphouët-Boigny, DFR Eaux, Forêts et Environnement, BP 1313, Yamoussoukro, Côte d'Ivoire
[223] Centre for Invasion Biology, Department of Mathematical Sciences, Stellenbosch University, Matieland 7602, South Africa
[224] African Institute for Mathematical Sciences, Muizenberg 7945, South Africa

from ~1.3 million sample plots. We then quantify drivers of local tree species richness patterns across latitudes. Generally, annual mean temperature was a dominant predictor of tree species richness, which is most consistent with the metabolic theory of biodiversity (MTB). However, MTB underestimated LDG in the tropics, where high species richness was moderated also by topographic, soil, and anthropogenic factors operating at local scales. Given that local landscape variables operate synergistically with bioclimatic factors in shaping the global LDG pattern, we suggest that MTB be extended to account for co-limitation by subordinate drivers.

Introduction

Identifying which mechanisms moderate global biodiversity patterns[1,2] has perplexed the scientific community for more than two centuries[3,4]. The most noticeable pattern, LDG, is a trend of declining local species richness (alpha diversity) from low to high latitudes. This trend has been observed for many taxonomic groups and across land, freshwater, and marine environments[5,6]. More than 30 hypotheses have been proposed[3,4,7,8] to explain LDG[9], but few can be reconciled with existing observational data for predicting biodiversity decline towards the poles. To test these varied hypotheses, biodiversity data must be assembled that are global in scope with sufficient sample coverage across all ecoregions and biomes.

In addition to biodiversity data, testing these varied hypotheses also requires data on a wide spectrum of potential drivers that may moderate biodiversity at local scales[9,10], such as climate, soil and land features, as well as anthropogenic factors. For instance, environmental temperature (*i.e.*, ambient temperature of the air, represented by annual mean temperature) is largely responsible for the generation and maintenance of biodiversity, through the effects of solar radiation on demographic rates (*e.g.*, growth and mortality), ecological interactions (*e.g.*, predation and competition) and evolutionary rates of change (*e.g.*, speciation and extinction)[11,12]. Soil and topographic heterogeneity facilitate niche partitioning via inducing microclimatic variation, contributing to compositional variation[13] and biodiversity maintenance[14,15]. Furthermore, humans have a long history of reshaping biodiversity through the selective use of natural resources and the modification of native species composition[16]. In addition, multiple subordinate factors jointly affecting biodiversity could potentially increase the diversity of niche opportunities, thereby resulting in species-rich assemblages.

Here, we quantified the relative contribution of a wide range of environmental factors across space on local tree species richness in forested areas around the world. To accomplish this, we standardized a global tree species richness (*i.e.*, as alpha diversity) database (**Fig. 1**) and quantified the relative contribution of 47 explanatory variables including bioclimatic conditions (*e.g.,* annual mean temperature), vegetation and survey attributes (*e.g.,* sample plot size), topographic covariates *(e.g.,* terrain roughness), soil covariates (*e.g.,* bulk density), and anthropogenic spatial features (*e.g.,* size of roadless areas) in an attempt to test whether local co-limitation exists when multiple subordinate drivers co-dominate (**Figs. 2&3**). We conducted a three-stage analysis (**Fig. 1**, see **Methods** in **Supplementary Information** for details) based on two independent ground-sourced forest inventory datasets (Phase-I and Phase-II, **Extended Data Fig. 1**). The main dataset (Phase-I) consisted of 1,255,444 sample plots, while the validating dataset (Phase-II) consisted of 22,131 sample plots, most of which are located in unsampled and under-sampled regions of the Phase-I dataset. Together, our sample data covered 424 of the 435 (97%) forested ecoregions worldwide (**Extended Data Figure 1**), with a total of ~55 million sample trees representing more than 32,000 species.

Results & Discussion

Global patterns of local tree species richness and latitudinal diversity gradient

Our analyses confirmed, with a high level of accuracy, one general spatial trend in local tree species richness worldwide that has led us to three conclusions regarding the mechanisms underlying patterns of tree species richness. We found that LDG for tree species richness was consistent with that of most other groups of organisms, with a decline from the tropics to the poles (**Figs. 2&4**). In the Northern Hemisphere, tree species richness dropped sharply from the equator (98 species·ha^{-1}) to 10°N with an average rate of decline of 6 species·ha^{-1} per 1° increase in latitude, after which the decline diminished and stabilized at 4 species·ha^{-1} at 50°N. In the Southern Hemisphere, tree species richness declined from the equator to 25°S on average by 3 species·ha^{-1} per 1° increase in latitude, after which tree species richness fluctuated before another steep drop from 25 species·ha^{-1} (43°S) to 4 species·ha^{-1} (50°S). We were able to detect and map regional patterns and global peaks of tree species diversity, with a high spatial resolution (0.025°×0.025°). The Amazonian, Southeast Asian, and Melanesian rainforests are clearly the regions with the greatest local tree species richness worldwide, containing >200 tree species·ha^{-1} above the 5 cm diameter-at-breast-height (DBH) threshold, confirming previous findings[17,18]. Tropical African rainforests generally contain 50% fewer tree species per hectare than Amazonian rainforests. In the temperate forests of the Northern Hemisphere, the Changbai Mountains in Northeast Asia (up to ~28 species·ha^{-1}) and the Central Appalachian forests in the Eastern United States (up to ~20 species·ha^{-1}) display high local species richness. In the Southern Hemisphere, the sclerophyllous and Nothofagus-dominated forests in south-central Chile are among the most species-rich temperate communities (up to 50 species·ha^{-1}). Boreal forest communities are consistently low in local tree species richness, with typically five or fewer tree species per hectare.

The above LDG pattern of tree species richness was generally consistent with the metabolic theory of biodiversity (MTB)[19,20], except at low latitudes (**Fig. 5**). According to MTB, environmental temperature is largely responsible for the generation and maintenance of biodiversity[12,21,22], and the natural logarithm of species richness is linearly associated with 1000/T, where T is the absolute environmental temperature in Kelvin (mean annual temperature + 273.15K), with a slope ranging from -7.5 to -9.0 K. Our global tree species richness gradient was largely consistent with MTB, with a slope of -8.0 K (p<0.001) and a coefficient of determination of 0.82 (see §***Metabolic Theory of Biodiversity*** in Supplementary Methods), indicating that environmental temperature is generally a good predictor of LDG. However, at low latitudes, MTB substantially underestimated LDG. In fact, near the equator where the actual LDG peaked (98 species·ha^{-1}), observed tree species richness was almost twice as high as predicted by MTB (56 species·ha^{-1}) (**Fig. 4A**). Our results suggest that within this low latitudinal range other factors are also important to the maintenance of biodiversity.

The under-estimation of local tree species richness by MTB at low latitudes is attributable, in part, to the lack of a definite dominant environmental factor, suggesting a **co-limitation** of multiple subordinate drivers at low latitudes (**Fig. 5**). In general, bioclimatic factors predominantly determined species richness in 82.6% of the forested areas, while co-limitation (*i.e.*, absence of any dominating factor) occurred in 11.7% of forested areas globally. However, in the low-latitude range between 5°N and 15°S, the percentage area of co-limitation increased to 37.1%, more than three times the global average. Furthermore, forested areas under co-limitation contained on average 81.1 ± 0.1 species per hectare, much higher than the average local tree species richness of forested areas predominantly determined by topographic (43.9 ± 0.1), anthropogenic (35.6 ± 0.2), soil (33.9 ± 0.2), and bioclimatic (19.4 ± 0.02) factors (**Fig. 5B**). This suggests that the pattern of co-limitation is pervasive in species-rich tropical forests. In South America, transitional areas between Amazonia and savanna formations nearby are subject to co-limitation that is partly attributable to a dynamic equilibrium between closed forest and savanna[23], edaphic conditions, and natural fire regimes[24]. In Africa, anthropogenic influences such as selective timber extraction and fuelwood collection, together with large-scale degradation[25] affect local tree species richness (**Fig. 5 & Extended Data Figure 7**). In Central

Africa, the evolution of anthropogenic influences from prehistoric to present times has imposed a substantial effect on species diversity[26] and resulted in the development of a complex system of mixes with light-demanding and old-growth tree species.

Bioclimatic dominance and co-limitation

In addition to an overall positive response of local tree species richness to the rise of annual mean temperature (see the partial dependence plot [PDP] of C_1 in **Fig.3** and **Extended Data Figure 3**), the importance of environmental temperature (2.7%) was topped by the total annual precipitation (C_{12}, 7.6%) (**Fig. 3**). Our findings are consistent with previous discoveries of a joint role of water and temperature/energy – as a proxy for net primary productivity[27] – on plant species richness, with water dominating particularly at warmer, lower latitudes[22,28]. Predicted tree species richness accelerated exponentially with temperature and rainfall, although independently, as shown in the cold-dry quadrant and the convex contours of the 2D PDP (**Extended Data Figure 3**), until each has reached its respective threshold (1500mm for total annual precipitation and 10°C for annual mean temperature). Beyond one of these thresholds, species richness is only limited by the predictor below its threshold (*i.e.*, by annual mean temperature in the cold-wet quadrant, or by annual precipitation in the hot-dry quadrant). When both predictors have reached their thresholds, *i.e.*, in the hot-wet quadrant, co-limitation predominates in most tropical forests. Net primary productivity in the tropics, thus, requires co-limitation of other factors besides only temperature and rainfall[29]. As the response of carbon flux mirror the low-latitude co-limitation pattern for tree species diversity, the matching determinants for both diversity and productivity may explain the similar latitudinal gradient in productivity, as well as and the positive diversity-productivity relationship[30,31]. Our findings also indicate that under climate change, intensified droughts coupled with increased annual mean temperature[32] can potentially trigger declines of tree species richness, although possible increases in water-use efficiency from elevated CO_2 and the dominance of highly contingent co-limiting factors may partially buffer this effect in the tropics[33].

Here, we articulate evidence for the *co-limitation* in LDG. Resource co-limitation is a common concept in ecology (e.g., 34,35), often used to describe how the synergistic interactions of two or more factors limit ecological productivity[36]. Our use of the term co-limitation emphasizes the reduced significance of a globally predominant driver of species richness at low latitudes, recognizing that several local subordinate factors synergistically contribute to increased tree species richness in this latitudinal range. We thus argue that the inclusion of co-limitation could substantially improve the explanatory power of biodiversity models in estimating alpha diversity by considering multiple subordinate factors where single-factor dominance is lacking, especially in the tropics. At high latitudes, bioclimatic conditions, particularly environmental temperature, are the major limiting factors, and thus the dominant drivers of tree species diversity. As the latitude declines, the influence of bioclimatic conditions dwindles, and the maintenance of tree species richness is moderated by many interacting drivers without a clear dominance, which is especially well expressed between 5°N and 15°S (**Fig. 5**). This prevalence of co-limiting factors is thus not a mere coincidence as to why the observed LDG at low latitudes is almost double that predicted by MTB (**Fig. 2**). While each of the existing hypotheses underpinning LDG addresses a certain process[10,12] (*e.g.*, selection, drift, dispersal, or speciation), the evidence of co-limitation highlights synergistic interactions of local processes across the latitudinal gradient.

Concluding remarks

More research is needed to fully elucidate patterns of LDG driven by climatic and other influences, especially those outlined in competing hypotheses. First, our analyses lack explicit consideration of some evolutionary, ecological and historical factors. These include mid-domain stochastic effects[37], the legacies of the poleward expansion of tree species after the Last Glacial Maximum[38,39], and recent human land use/management. Alternative hypotheses, such as niche conservatism or climatic history, are more difficult to test due to data limitations. In addition, long-term effects at geological and millennial time scales also play a role, but it is difficult to disentangle these effects due to collinearity[40]. A major source of uncertainty in our results (**Fig. 4B**) came from an uneven sample coverage between developed and developing countries (**Extended Data Fig. 1**). To address this gap, we argue that there needs to be a shared responsibility among forestry agencies at various levels of government, scientists, indigenous communities, and other biodiversity monitoring groups to improve sample coverage of forest inventories in developing countries. Innovative biodiversity funding mechanisms, *e.g.*, forest inventories funded by carbon initiatives such as REDD+, should be incorporated into a comprehensive global forest biodiversity database. Meanwhile, the severe shortage of experts and database management infrastructures, especially in developing countries, poses another major challenge to address this gap[41]. The education and training of new generations of forest scientists, taxonomists, and foresters can bring tangible benefits to biodiversity monitoring while improving local economies as well.

Considering co-limitation in addition to MTB enables a refined description of the biogeographic distribution of biodiversity and mechanisms underlying LDG. Our analysis has resulted in the production of a high-resolution map of tree species richness across the global forest range, along with visuals of those factors responsible for the moderation of local tree species richness. Such tools are necessary for conservation management which requires assessments of factors responsible for biodiversity patterns at multiple scales that matter – from local, regional to global scales. Patterns of local tree species richness and associated drivers may provide insights into how and why the diversity of other forest flora, fauna, and microbes[42,43] vary across space and time. Furthermore, the high-resolution map of local tree species richness presented here provides a benchmark for evaluating the impact of biodiversity loss on the productivity and functioning of forest ecosystems[31,44]. Finally, aligned with current international calls for spatially explicit monitoring of ecosystem attributes[45], this study delivers detailed biogeographic information to support international endeavors[46] focused on valuing natural capital and advancing global conservation.

METHODS

As illustrated in **Fig. 1**, we conducted data analyses and modeling in three stages.

Stage 1: Data Standardization

For this study, we compiled individual *in situ* tree data from all the regional and national GFBi forest inventory data sets (**Table S2**) into a standardized GFBi dataframe, *i.e.* the GFBi tree list. In this standardized GFBi dataframe, each row represents an individual tree, and columns represent nine key tree- and plot-level attributes. These attributes are tree ID (FID), a unique number assigned to each individual tree; plot ID (PLT), a unique string assigned to each plot; plot coordinates (LAT and LON); tree species name (SPCD); diameter-at-breast-height (DBH) or above buttress; year of measurement; and data set name (DSN), a unique number assigned to each forest inventory data set (**Table S2**). With a total of 56 million trees surveyed, GFBi individual-based dataframe represents 1/50,000 of the approximately 2.7 trillion trees[47] worldwide. Because all trees in each sample plot were identified and measured, GFBi data make it possible to quantify forest community structure, composition, and species distribution.

To ensure consistency and maximum accuracy in species names, we standardized observations from different forest inventory data sets with the following protocol. First, all multi-stem trees were divided so that each stem represents an individual tree. The scientific names were extracted from original data sets, keeping only the genus and species (authority names were removed). Next, all the species names were compiled into five general species lists, one for each continent. We verified individual species names against 23 online taxonomic databases or web application programming interfaces (API) using the gnr_resolve() function from the 'taxize' package[48] of R[49]. We then manually verified and corrected all the names that did not match with the majority of the online taxonomic databases, that is, the names with a matching score lower than 0.9. For individuals denoted by morphospecies, we assigned each a unique name comprising the genus name and a unique species code. The unique species code consisted of the string "spp", plus the data set name followed by a unique number denoting if two individuals belong to the same species. For example, "*Aidia* sppCDi1" and "*Aidia* sppCDi2" represented two different species under the genus "*Aidia*", and both species have been observed in a forest inventory of the Democratic Republic of the Congo named "CDi." To maximize our species coverage, a tree was defined in this study as a perennial plant with an elongated woody stem that supports branches and leaves, including woody angiosperms, gymnosperms, and taller palms (Arecaceae). Tree ferns (Cyatheales) and bamboos (Bambusoideae) were excluded from our analysis.

From the GFBi individual-tree-level dataframe, we derived a ***global species abundance matrix*** (GSAM). The GSAM consisted of the number of individuals by species (column vectors) within individual sample plots (row vectors). The global species abundance matrix consisted of two complementary datasets: Phase-I dataset contained 1,255,444 sample plots, and Phase-II dataset contained 22,131 sample plots, most of which are located in unsampled and under-sampled regions of Phase-I dataset. Phase-I sample plots cover 394 ecoregions across the world, and Phase-II sample plots cover an additional 30 ecoregions in Africa, South America, Southeast Asia, Mexico, India, and Japan. Together, our ground-based forest sample plots cover 424 of 435 (97.5%) forested ecoregions across the world. The global species abundance matrix contains ~1.3 million rows (plots) by 32,608 columns (species). Key plot-level information was added to the matrix, including PLT, DSN, plot coordinates, basal area (B), the total cross-sectional areas (m^2) of living trees per hectare calculated from DBH and TPH (expansion factor), and the year of measurement. TPH denotes the number of trees per hectare represented by each sampled individual. It ranged from 1 to 5,244 across the GFBi data, with a mean of 48 trees per hectare.

We quantified, for each sample plot, tree species richness (*S*) which is the total number of tree species in a community. Due to the difference in plot size (SD=0.09 ha) and threshold DBH values (SD=2.52 cm) across GFBi sample plots, we developed machine learning models to standardize tree species richness for a common basis of 1 ha in area and 5 cm in threshold DBH. The models incorporated both plot area (*A*) and threshold DBH (*D*) as predictors to account for the underlying species-area relationship[50-52] and species-individual size distribution[53] in a rarefaction-based approach[54]. This standardization approach justifies compiling direct tree species diversity estimates from GFBi *in situ* data of different sources and sampling protocols[55-57], an issue highlighted in earlier large scale — although significantly less extensive — forest biodiversity studies[57,58]. To evaluate the accuracy of this standardization approach, we tested the machine learning models using cross-sample validation, and compared our global maps of estimated tree species diversity against other standardization approaches based on sample completeness (see §**Model Evaluation** below).

The machine learning models employed 47 environmental covariates to predict tree species richness. These covariates, derived from satellite-based remote sensing and ground-based survey data, can be summarized into five general categories: *bioclimatic* (*e.g.*, annual mean temperature, total annual precipitation, potential evapotranspiration, and indexed annual aridity); *soil* (bulk density, pH, electrical conductivity, C/N ratio, and total nitrogen); *topographic*, including elevation, slope, aspect, and terrain features; *vegetation and survey* attributes (plot size, basal area, threshold diameter, and percent forest canopy cover); and *anthropogenic* variables (human footprint, roadless areas, and size of protected areas) (**Table S1**). We extracted all geospatial covariate values from raster datasets to point locations of GFBi plots using ArcMap 10.3[59] and R 3.4.1[49], to build a ***standardized plot-level dataframe***.

Stage 2: Model Training and Evaluation

We trained random forests (RF)[60], an ensemble learning method that detects general trends present in the data using a multitude of decision trees, to estimate standardized community-level tree species diversity. The RF algorithm applies the general technique of bootstrap aggregating (bagging) with a modified tree learning algorithm that selects, at each candidate split in the learning process, a random subset of the features (*i.e.*, feature bagging). Since a random subset of variables is chosen for each tree, the RF algorithm based on bagged tree ensembles avoids overfitting[60] and mitigates the multicollinearity issue[61] posed by high correlations between some of the predictors variables (**Fig. 3**). Using subsamples of GFBi data as the training set (*i.e.*, ***training dataframe***) with response S, bagging repeatedly for B times selects a random sample with replacement of the training set and trains a regression tree f_b. After training, RF can predict for unseen samples $\mathbf{X'}$, with the response variable S being tree species richness per hectare:

$$S = \frac{1}{B}\sum_{b=1}^{B} f_b(\mathbf{X}'). \quad (1)$$

For rigorous model evaluation, we employed three very different cross-validation approaches: randomized cross-validation (RCV), spatial cross-validation (SCV), and post-sample validation (PSV). In randomized cross-validation (RCV), a model was trained for each continent with a random subsample that accounted for 90% of the training data from that continent, and the remaining 10% of the training data were used as the testing set. This process was repeated 20 times with sample replacement to examine the accuracy of estimated tree species diversity values. In spatial cross-validation (SCV), all sample data from an ecoregion[62] were reserved for testing the model that was trained with the remaining samples from the larger continent within which the ecoregion is situated. We decided to use ecoregions as spatial blocks because 1) unlike political units such as countries and provinces, ecoregions are delineated based on ecological and bioclimatic conditions; and 2) with a total of ~700 terrestrial ecoregions across the world, each ecoregion encompasses 1,800 sample plots on average, which is a large enough sample size for training RF models. This process was repeated until all the forested ecoregions across the world had been tested. SCV was more rigorous than RCV, because samples from an entire ecoregion rather than random samples were withheld for validation. Post-sample validation (PSV) was the most rigorous among the three validation processes. For PSV, we have collated an independent sample dataset from 22,131 forest sample plots, which we named Phase-II sample plots to highlight their independence from the original GFBi dataset (*i.e.*, Phase-I sample plots). In PSV, we used Phase-II data as the testing set to evaluate the accuracy of the predictive models that were trained for each continent with the Phase-I data.

Using these three cross-validation processes, we also evaluated the performance of the RF model against two other predictive models, including multiple regression with ordinary least squares (OLS) and Extreme Gradient Boosting (XGBoost). For each model, we derived predicted values of tree species richness of the testing sets, and compared these predicted values against observed data using mean absolute error (MAE), root-mean-squared error (RMSE), and coefficient of determination (R^2)[63]. The process was repeated 20 times to select the best model for each continent.

The OLS model estimated values of standardized point diversity for non-sampled point location $\mathbf{s}$, based on spatially explicit values of covariates:

$$Y(\mathbf{s}) = \boldsymbol{\alpha}\cdot\mathbf{X}(\mathbf{s}) + \mathbf{e}(\mathbf{s}), \qquad (2)$$

where $Y(\mathbf{s})$ is tree species richness at location $\mathbf{s}$; $\mathbf{X}$ a design matrix for the predictor variables at location $\mathbf{s}$; $\boldsymbol{\alpha}$ is a vector of coefficients; and $\mathbf{e}$ is a random vector following a Gaussian probability density function, with an expected value of zero and variance of $\boldsymbol{\sigma}^2$. Spatial autocorrelation[64] was not accounted for here due to computational limitations. GFBi data collected from sample plots of various sizes were harmonized to represent local forest community populations per hectare using the expansion factor[65], and we used the standardized species richness per hectare values, for the response variables. We fit a model (2) for each continent. To mitigate the multicollinearity issue[66], we selected for the OLS model the best subset of predictor variables for each continent from the predictor variables used in the RF models, using step-wise regression and Akaike information criterion[67].

XGBoost is a scalable machine learning system[68] that implements the gradient boosting decision tree algorithm[69]. With this ensemble technique, an initial model was trained, with new models added sequentially to correct for errors made by each existing model until no further improvements could be made. Then, new and initial models were merged to make a final prediction that minimized errors. With its algorithm engineered for efficiency in computing time and memory resources, XGBoost is widely used by data scientists to achieve state-of-the-art results on a number of machine learning challenges[68]. In this study, the XGBoost model estimated tree species diversity values in three steps. First, an initial model F_0 was defined to predict the target variable Y. This model was associated with a residual $(Y - F_0)$. Second, a new model h_1 was fit to the residuals from the previous step, and F_0 and h_1 were combined to form the boosted model F_1:

$$F_1(\mathbf{x}) = F_0(\mathbf{x}) + h_1(\mathbf{x}), \tag{3}$$

of which the mean squared error was lower than that from F_0. Finally, to improve the performance of F_1, we modelled after the residuals of F_1 to create a new model F_2, and repeated it for m iterations, until the mean squared error converged:

$$F_m(\mathbf{x}) = F_{m-1}(\mathbf{x}) + h_m(\mathbf{x}). \tag{4}$$

Before training RF and XGBoost models, we fine-tuned four key hyper-parameters, two for each model. Using 20 bootstrapping iterations on random training sets consisting of 90% of the samples, we first evaluated the sensitivity of RMSE of the testing sets (consisting of the remaining 10% of the samples) to the number of trees to grow and the number of variables randomly sampled as candidates at each split for the RF model, and selected the optimal hyper-parameter values (**Extended Data Figure 5**). Similarly, we selected the optimal values of the maximum number of boosting iterations (*i.e.*, number of rounds), and the maximum depth of a tree for the XGBoost model (**Extended Data Figure 6**). As a result, we obtained a ***preliminary RF model***.

Because the RF model emerged as the most accurate model from all three cross-validation processes (**Extended Data Figure 2**), we selected the RF as the final model, and re-calibrated the ***final RF model*** using all the sample data (Phase-I and Phase-II data).

Stage 3: Global Species Richness Assessments

Global map of local tree species richness

To map community-level tree species richness over the global forest range, we first derived the global forest range map from version 1.3 of the Global Forest Change database[70] (years 2000-2015). To ensure consistency with the United Nations FAO's definition of forest[71], the global forest range in this study was defined as forested areas with ≥ 10% tree crown coverage per unit area. The tiled 'treecover2000', 'loss', and 'gain' datasets were integrated to obtain current forest cover estimates for the year 2015. To minimize processing artefacts, the ~1 arc-second spatial resolution tiles were spatially aggregated to an even multiple of their native resolution that approximated the resolution of our covariates. The datasets were then converted to vector point files before being reconverted to raster format with the exact resolution and origin of our covariates. After mosaicking each set of tiles, we computed 'treecover' (scaled) – 'loss' + 'gain' to obtain the 2015 global forest cover, represented as percent forest cover per ~30 arc-second pixel. Artefacts in the original data led to 0.08% of all terrestrial pixels having forest cover estimates greater than 100% and 1.9% of terrestrial pixels having estimates less than 0%. These values were truncated to 100% and inflated to 0%, respectively. Finally, the global forest range consisted of those pixels with a percent forest cover ≥10% in 2015. In total, each map consisted of 9,944,908 pixels of 0.025° × 0.025° (hereafter, the pixel) of forested areas. This range is rather conservative and potentially underestimates many remnant forests in drylands and grasslands[72].

We then estimated *tree species richness* at a one-hectare scale for all pixels within a continent based on the final RF model trained for that continent, using both Phase-I and Phase-II data. Spatially explicit local environmental covariate data across the global forest range were used for the imputation, except that plot size and threshold diameter-at-breast-height were set as 1 ha and 5 cm, respectively. For ecoregions with extremely low sample coverage, we further fine-tuned the RF model using samples of similar environment characteristics from other continents. More specifically, we first identified two ecoregions of extremely low sample coverage, *i.e.*, the temperate forests in South America and the tropical forests in Oceania, as there were fewer than 1000 sample plots for the entire biome on those continents. We then trained a new RF model for each ecoregion, using all the sample data from the same biome across the world, and fine-tuned the mapping data for that ecoregion using the biome-specific RF model.

We computed and mapped the width of the *95% confidence interval* for our local estimates of tree species richness per hectare across the global forest range. To this end, we employed a rigorous spatial-block approach, analogous to the spatial cross-validation, to derive the 95% confidence interval. More specifically, we computed the width of the 95% confidence interval for each 0.025° × 0.025° mapping pixel by ecoregion. For a pixel p in ecoregion e, we trained 20 RF models using random subsamples that accounted for 90% of the training data from the same continent, which included all samples except those from ecoregion e. We then derived the standard error and the width of the 95% confidence interval for this pixel p in ecoregion e, from the predictions of the 20 RF models trained for this ecoregion. This process was repeated until all the forested ecoregions across the world had been assessed and mapped.

Uncertainty in our global diversity estimates was caused by two types of errors. The first was measurement error from *in situ* forest inventories. We mitigated this type of error by implementing stringent species-name check and data standardization protocols (see §**Stage 1 Data Standardization**). The second arose from the imputation process to map tree species diversity. We minimized this type of error using the three cross-validation approaches introduced in **§Stage 2**.

Metabolic theory of biodiversity

Using the global standardized tree species richness values predicted from the final RF models, we quantified the ***global latitudinal diversity gradient*** (LDG) of tree species richness, and tested the effect of environmental temperature based on the metabolic theory of biodiversity (MTB)[19]:

$$\ln(S) = \alpha \frac{1000}{T_{env}} + \beta, \tag{5}$$

where S represents species richness, and T_{env} here represents absolute environmental temperature (mean annual temperature + 273.15K); α and β represent coefficients to be estimated by ordinary least squares. According to original and extended MTB[19,20], the slope α is expected to range between -7.5 and -9.0 Kelvin, under the assumption that tree community abundance per-area does not vary with latitude.

Variance Partitioning

We used variance partitioning[73], based on the sample data from ~1.3 million plots, to quantify the unique and joint fractions of spatial variance in tree species richness explained by environmental factors and latitude. Due to the correlation between species and environment, and between the spatially explicit environmental factors, the variance partitioning approach mitigates type-I error inflated by spatial autocorrelation[74]. With variance partitioning, we tested the significance of environmental effects on tree species richness in a series of nested Random Forest (RF) models. (A) *The full model* (**Extended Data Figure 4A**) consisted of latitude and 47 environmental variables (including 21 bioclimatic ones). (B) *The reduced model I* (**Extended**

Data Figure 4B) consisted of all but the 21 bioclimatic variables. (C) *The reduced model II* (**Extended Data Figure 4C**) consisted of only a zero constant. The overall significance of all environmental factors plus latitude was tested in an one-tailed F-test by comparing the residual sum of squares of error (RSS) of model (A) and model (C):

$$F = \frac{\frac{RSS_C - RSS_A}{n_C - n_A}}{\frac{RSS_A}{n - n_A}}, \qquad (6)$$

where $n - n_A$ and $n_C - n_A$ stand for the degree of freedom for the full model, and the difference in the degrees of freedom between the full model and the reduced model II, respectively.

The significance of bioclimatic factors, with the effect of latitude being controlled, was tested in an one-tailed F-test by comparing RSS of model (A) and model (B):

$$F = \frac{\frac{RSS_B - RSS_A}{n_B - n_A}}{\frac{RSS_A}{n - n_A}}, \qquad (7)$$

where $n_B - n_A = 21$ stands for the difference in the degrees of freedom between the full model and the reduced model I.

We partitioned the spatial variance in observed species richness into four components: [*a*] represents the fraction of variance uniquely explained by environmental factors (*i.e.*, bioclimatic, topographic, anthropogenic, and soil variables), after latitudinal effects have been taken into account; [*b*] represents the fraction of variance jointly explained by environmental factors and latitudinal effects; [*c*] represents the fraction of variance explained by latitudinal effects after removing environmental effects; and [*d*] represents the fraction of variance not explained by the full RF model. Then, the total fraction of variance explained by both environmental factors and latitude was [*a* + *b* + *c*], the fraction of variance explained by environmental factors was [*a* + *b*], and the fraction of variance explained by latitude was [*b* + *c*]. Components [*a* + *b* + *c*], [*a* + *b*] and [*b* + *c*] were estimated by the R^2 statistics from the RF models trained for each continent using all factors, environmental factors, and latitude, respectively (see §**Stage 2 Model Training and Evaluation**). Components [*a*], [*b*] and [*c*] were computed from the previous components using arithmetic relationships that ensure that [*a*] + [*b*] + [*c*] + [*d*] = 100%.

Model sensitivity

Based on the final RF models and sample data from ~1.3 million plots, we mapped the dominant drivers of tree species richness with a 0.025° × 0.025° resolution (*i.e.*, ***global map of co-limitation***), following a standard procedure for model sensitivity analysis[75]:

Step 1: Using the full RF model, and the values of environmental factors $\mathbf{X(s)}$ specific to a 0.025°-pixel **s**, we had already estimated local tree species richness $S_{full}(\mathbf{s})$:

$$S_{full}(\mathbf{s}) = f\big(\mathbf{X}(\mathbf{s})\big), \tag{8}$$

where *f*() represents the RF model, and $\mathbf{X(s)}$ environmental factors in four categories, namely **E1**: bioclimatic, **E2**: topographic, **E3**: anthropogenic, and **E4**: soil.

Step 2: For the above-mentioned pixel, we estimated a new local tree species richness value $S_{-E1}(\mathbf{s})$, using a reduced RF model in which all **E1** (bioclimatic) variables were removed:

$$S_{-E1}(\mathbf{s}) = f_{-E1}\big((\mathbf{X}\text{-}\mathbf{E1})(\mathbf{s})\big), \tag{9}$$

where $f_{-E1}()$ represents the RF model trained with all but 21 bioclimatic variables, and $(\mathbf{X}\text{-}\mathbf{E1})(\mathbf{s})$ encompassed environmental factors in three categories, namely **E2**: topographic, **E3**: anthropogenic, and **E4**: soil.

Step 3: For a given pixel, we calculated the relative sensitivity of predicted species richness to **E1**:

$$R(E1) = \left|S_{full}(\mathbf{s}) - S_{-E1}(\mathbf{s})\right| / S_{full}(\mathbf{s}) \quad . \tag{10}$$

Step 4: We repeated Steps #2 and #3 to calculate, for a given pixel, the relative sensitivity of each of the following categories (*i.e.*, **E2**: topographic, **E3**: anthropogenic, and **E4**: soil), respectively. The dominant driver (*i.e.*, limiting factor) for this pixel was then the category with the highest relative sensitivity, provided that this relative sensitivity was greater than or equal to 0.2.

Step 5: If the relative sensitivities were less than 0.2 for all categories, we considered that this was a scenario of joint effects of multiple categories of factors (*i.e.*, co-limitation), rather than dominance of a single category. Where clear dominance of a single category was lacking, we denoted the dominant driver of this pixel as ‘**E5**: co-limitation.’

Step 6: We repeated the steps above to calculate, for all the remaining pixels of the global grid, the relative sensitivity of each of the five categories of environmental factors, namely **E1**: bioclimatic, **E2**: topographic, **E3**: anthropogenic, **E4**: soil, and **E5**: co-limitation. Based on these values, we created a wall-to-wall map of dominant drivers of tree species richness across the global forest range, by labeling the category with the highest relative sensitivity for each pixel (**Fig. 5A**).

Step 7: Based on the relative sensitivity obtained from the Steps #1-6, we computed percent prevalence (0–100%) of bioclimatic, topographic, anthropogenic, and soil factors, as well as a lack of dominance (co-limitation) in all the forested pixels along each latitudinal band.

Data availability

- The global map of tree species richness is available under license CC BY 4.0, with the identifier: 10.6084/m9.figshare.17232491. This map can be downloaded in two formats. One is a geoTIFF file (S_mean_raster.tif) containing the fully geo-referenced map of tree species richness worldwide at a 0.025°×0.025° resolution. The other is a comma-separated file (S_mean_grid.csv) with the following attributes:

  S is local average tree species richness per hectare

  x, y are centroid coordinates of all 0.025°×0.025° pixels;

- The global map of co-limitation is available under license CC BY 4.0, with the identifier: 10.6084/m9.figshare.17234339.

- The metadata of the entire training dataframe – including the characteristics and references of all the *in situ* Phase-I and Phase-II datasets, as well as the definitions, units, and summary statistics of the environmental covariates – is available under license CC BY 4.0, with the identifier: 10.6084/m9.figshare.19733449.v1

- The public version of the training dataframe including the plot-level species richness and all the covariates, which is needed to reproduce the models and results presented here, is available at: https://doi.org/10.6084/m9.figshare.20055488. The dataframe is also available on two international web research platforms: science-i.org, and gfbinitiative.org.

- Raw forest inventory data are commonly subject to a wide array of confidentiality clauses in regard to open access policies. Despite recent efforts to make some of these data fully open[76,77], some governments and private data owners, especially those from the developing countries generally have decided to keep their data confidential. This decision is based on well-founded arguments to protect certain trees or forests (because of their large size or protected taxonomic status) from illegal logging or trespassing, and to protect landowners' privacy, against the misuse of plot information such as the geographic coordinates. The sensitive information in the training dataframe, including the plot coordinates and tree-level information, will be available from the corresponding author (albeca.liang@gmail.com) upon a request via Science-I or GFBI, and an approval from data contributors.

Code availability

All the models in this study were constructed using command line applications written in the R programming language, which processed and restructured the input data, trained the model, and performed cross-validation. Due to the massive amount of data, we used Purdue University's Brown supercomputing cluster to accelerate the training process.

The development of the GFBi database, tabular data cleaning, creation of species abundance matrices, evaluation of diversity determinants, and geostatistical imputation were conducted in R[49] (v.3.4.2) through the use of several Linux-based high-performance computing

(HPC) resources at Purdue University, and a custom HPC interface developed using Amazon Web Services, each designed for batch processing, scalable resource distribution, embarrassingly parallel computations, and/or large RAM jobs. Compute nodes with up to 1TB of RAM and clusters of up to 64 nodes were employed in this study. Portions of the covariate preparation, mapping, and quality control assessment were conducted on Windows-based operating systems with up to 128 GB of RAM.

Final continental-level RF models and the R codes we developed to train the models are available under license MIT, with the identifier: 10.6084/m9.figshare.17234729.

Inclusion & Ethics statement

The international research collaboration leading to this research paper was conducted via Science-i.org, a transparent and FAIR (Findable, Accessible, Interoperable, and Reusable) web platform for international research collaboration. Through this platform and our partner initiatives including the Global Forest Biodiversity Initiative (GFBI), we pursue excellence and high standards of performance, professionalism, and ethical conduct. Science-i strictly prohibits any form of discrimination against individual on the basis of gender, race, age, religion, sexual orientation, veteran status, or disability status. Science-i continuously seek and encourage underrepresented and underprivileged people and groups, as well as the unique voices in global scientific research collaboration.

ACKNOWLEDGMENTS

The team collaboration and manuscript development are supported by the web-based team science platform: science-i.org, with the project #202205GFB2. We thank the following initiatives, agencies, teams, and individuals for data collection and other technical support: the Global Forest Biodiversity Initiative (GFBI) for establishing the data standards and collaborative framework; United States Department of Agriculture, Forest Service, Forest Inventory and Analysis (FIA) Program; University of Alaska Fairbanks; The SODEFOR, Ivory Coast; University Félix Houphouët-Boigny (UFHB, Ivory Coast); the Queensland Herbarium and past Queensland Government Forestry and Natural Resource Management Departments and staff for data collection for over seven decades; the National Forestry Commission of Mexico (CONAFOR). We thank Marc Baker (Carbon Tanzania), together with a team of field assistants (Valentine and Lawrence); all persons who made the Third Spanish Forest Inventory possible, especially the main coordinator, J. A. Villanueva (IFN3); The French National Forest Inventory (NFI campaigns [raw data 2005 and following annual surveys, were downloaded by GFBI at https://inventaire-forestier.ign.fr/spip.php?rubrique159, site accessed on January 1st, 2015); the Italian Forest Inventory (NFI campaigns Raw data 2005 and following surveys were downloaded by GFBI at https://inventarioforestale.org/, site accessed on April 27th, 2019); Swiss National Forest Inventory, Swiss Federal Institute for Forest, Snow and Landscape Research WSL and Federal Office for the Environment FOEN, Switzerland; The Swedish NFI, Department of Forest Resource Management, Swedish University of Agricultural Sciences SLU; The National Research Foundation (NRF) of South Africa (89967 and 109244) and the South African Research Chair Initiative; The Danish National Forestry, Department of Geosciences and Natural Resource Management, UCPH; Coordination for the Improvement of Higher Education Personnel of Brazil (CAPES, grant no. 88881.064976/2014-01); Rafael Ávila and Sharon van Tuylen, Instituto Nacional de Bosques (INAB), Guatemala for facilitating Guatemalan data; The National Focal Center for Forest condition monitoring of Serbia (NFC), Institute of Forestry, Belgrade, Serbia; The Thünen Institute of Forest Ecosystems (Germany) for providing National Forest Inventory data; the Food and Agriculture Organization of the United Nations (FAO) and the United Nations High Commissioner for Refugees (UNHCR) undertaking the SAFE (Safe Access to Fuel and Energy) and CBIT-Forest projects; the Amazon Forest Inventory Network (RAINFOR), the African Tropical Rainforest Observation Network (AfriTRON), and the ForestPlots.net initiative for their contributions from Amazonian and African forests. The Natural Forest plot data collected between January 2009 and March 2014 by the LUCAS programme for the New Zealand Ministry for the Environment, are provided by the New Zealand National Vegetation Survey Databank https://nvs.landcareresearch.co.nz/. The International Boreal Forest Research Association (IBFRA). The Forestry Corporation of New South Wales, Australia. The National Forest Directory of the Ministry of Environment and Sustainable Development of the Argentine Republic (MAyDS) for the plot data of the Second National Forest Inventory (INBN2). The National Forestry Authority of Uganda for their National Biomass Survey (NBS) dataset. The Sabah Biodiversity Council and the staff from Sabah Forest Research Centre. All TEAM data are provided by the Tropical Ecology Assessment and Monitoring (TEAM) Network, a collaboration between Conservation International, the Missouri Botanical Garden, the Smithsonian Institution, and the Wildlife Conservation Society, and partially funded by these institutions, the Gordon and Betty Moore Foundation, and other donors, with thanks to all current and previous TEAM site manager and other collaborators that helped collecting data; The people of the Redidoti, Pierrekondre and Cassipora village who were instrumental in assisting with the collection of data and sharing local knowledge of their forest; and the dedicated members of the field crew of Kabo 2012 census; the National Forestry Authority and Ministry of Water and Environment of Uganda for their National Biomass Survey (NBS) dataset. This research was supported in part through computational resources provided by Information Technology at Purdue, West Lafayette, Indiana.

Funding

This work is supported in part by the NASA Grant #12000401 “Multi-sensor biodiversity framework developed from bioacoustic and space based sensor platforms” (JL, BP); the USDA National Institute of Food and Agriculture McIntire Stennis projects 1017711 (JL) and 1016676 (MZ); the US National Science Foundation Biological Integration Institutes grant NSF-DBI-2021898 (PBR); the funding by H2020 VERIFY (contract 776810) and H2020 Resonate (contract 101000574) (GJN); The TEAM project in Uganda supported by the Moore foundation and Buffett Foundation through Conservation International (CI) and Wildlife Conservation Society (WCS); The Danish Council for Independent Research | Natural Sciences (TREECHANGE, grant 6108-00078B) and VILLUM FONDEN grant #16549 (JCS); The Natural Environment Research Council of the UK (NERC) project NE/T011084/1 awarded to JAG and NE/ S011811/1; ERC Advanced Grant 291585 (“T-FORCES”) and a Royal Society-Wolfson Research Merit Award (OLP); RAINFOR plots supported by the Gordon and Betty Moore Foundation and the U.K. Natural Environment Research Council, notably NERC Consortium Grants ‘AMAZONICA’ (NE/F005806/1), ‘TROBIT’ (NE/D005590/1), and ‘BIO-RED’ (NE/N012542/1); CIFOR's Global Comparative Study on REDD+ funded by the Norwegian Agency for Development Cooperation, the Australian Department of Foreign Affairs and Trade, the European Union, the International Climate Initiative (IKI) of the German Federal Ministry for the Environment, Nature Conservation, Building and Nuclear Safety, and the CGIAR Research Program on Forests, Trees and Agroforestry (CRP-FTA), and donors to the CGIAR Fund; AfriTRON network plots funded by the local communities and NERC, ERC, European Union, Royal Society and Leverhume Trust; A grant from the Royal Society and the Natural Environment Research Council, UK (SLL); National Science Foundation CIF21 DIBBs: EI: #1724728 (ACC); National Natural Science Foundation of China (31800374) and Shandong Provincial Natural Science Foundation (ZR2019BC083) (HL). UK NERC Independent Research Fellowship (grant code: NE/S01537X/1) (TJ); a Serra-Húnter Fellowship provided by the Government of Catalonia (Spain) (S. d-M); the Brazilian National Council for Scientific and Technological Development (CNPq, grant 442640/2018-8, CNPq/Prevfogo-Ibama N◦ 33/2018) (CAS); a grant from the Franklinia Foundation (DAC); Russian Science Foundation project

Project # 19-77-300-12 (RV); the Takenaka Scholarship Foundation (AOA); the German Research Foundation (DFG), grant number Am 149/16-4 (CA); the Romania National Council for Higher Education Funding, CNFIS, project number CNFIS-FDI-2022-0259 (OB); Natural Sciences and Engineering Research Council of Canada [RGPIN-2019-05109 and STPGP506284] and the Canadian Foundation for Innovation (36014) (HYHC); the project SustES – Adaptation strategies for sustainable ecosystem services and food security under adverse environmental conditions (CZ.02.1.01/0.0/0.0/16_019/0000797) (EC); Consejo de Ciencia y Tecnología del estado de Durango (2019-01-155) (JJC-R); Science and Engineering Research Board (SERB), New Delhi, Govt. of India (File No. PDF/2015/000447) - "*Assessing the carbon sequestration potential of different forest types in Central India in response to climate change*" (JAD); Investissement d’avenir grant of the ANR [CEBA: ANR-10-LABEX-0025] (GD); National Foundation for Science & Technology Development of Vietnam, 106-NN.06-2013.01 (TVD); Queensland government, Department of Environment and Science (TJE); A Czech Science Foundation Standard Grant (19-14620S) (TMF); European Union Seventh Framework Program (FP7/2007-2013) under grant agreement n° 265171 (LF); Grants from the Swedish National Forest Inventory, Swedish University of Agricultural Sciences (JF); CNPq productivity grant nº 311303/2020-0 (ALdG); DFG grant HE 2719/11-1,2,3; HE 2719/14-1 (AH); European Union's Horizon Europe research project OpenEarthMonitor grant number 101059548, CGIAR Fund INIT-32-MItigation and Transformation Initiative for GHG reductions of Agrifood systems RelaTed Emissions (MITIGATE+) (MH); General Directorate of the State Forests, Poland (1/07; OR-2717/3/11; OR.271.3.3.2017) and the National Centre for Research and Development, Poland (BIOSTRATEG1/267755/4/NCBR/2015) (AMJ); Czech Science Foundation 18-10781S (SJ); Danish of Ministry of Environment, The Danish Environmental Protection Agency, Integrated Forest Monitoring Program – NFI (VKJ); State of São Paulo Research Foundation/FAPESP as part of the BIOTA/FAPESP Program Project Functional Gradient-PELD/BIOTA-ECOFOR 2003/12595-7 & 2012/51872-5 (CAJ); Danish Council for Independent Research - social sciences - grant DFF 6109–00296 (GAK); Russian Science Foundation project 21-46-07002 for the plot data collected in the Krasnoyarsk region (VK); BOLFOR (DKK); Department of Biotechnology, New Delhi, Govt. of India (Grant No. BT/PR7928/NDB/52/9/2006, dated 29 September 2006) (MLK); Grant from Kenya Coastal Development Project (KCDP), which was funded by World Bank (JNK); Korea Forest Service (2018113A00-1820-BB01, 2013069A00-1819-AA03, and 2020185D10-2022-AA02), and Seoul National University Big Data Institute through the Data Science Research Project 2016 (HSK); The Brazilian National Council for Scientific and Technological Development (CNPq, grant 442640/2018-8, CNPq/Prevfogo-Ibama N◦ 33/2018) (CK); CSIR, New Delhi, Govt. of India (Grant No. 38(1318)12/EMR-II, Dated: 03.04.2012) (SK); Department of Biotechnology, New Delhi, Govt. of India (Grant No. BT/ PR12899/ NDB/39/506/2015 dated 20/06/2017) (AK); Coordination for the Improvement of Higher Education Personnel (CAPES) #88887.463733/2019-00 (RVL); National Natural Science Foundation of China (31800374) (HL); Project of CEPF RAS "Methodological approaches to assessing the structural organization and functioning of forest ecosystems" (AAAA-A18-118052590019-7) funded by the Ministry of Science and Higher Education of Russia (NVL); Leverhulme Trust grant to Andrew Balmford, Simon Lewis and Jon Lovett (ARM); Russian Science Foundation, project 19-77-30015 for European Russia data processing (OM); Grant from Kenya Coastal Development Project (KCDP), which was funded by World Bank (MTEB); The National Centre for Research and Development, Poland (BIOSTRATEG1/267755/4/NCBR/2015) (SM); the Secretariat for Universities and of the Ministry of Business and Knowledge of the Government of Catalonia and the European Social Fund (AM); Queensland government, Department of Environment and Science (VJN); Pinnacle Group Cameroon PLC (LNN); Queensland government, Department of Environment and Science (MRN); the Natural Sciences and Engineering Research Council of Canada (RGPIN-2018-05201) (AP); The Russian Foundation for Basic Research, project # 20-05-00540 (EIP); European Union Seventh Framework Program (FP7/2007-2013) under grant agreement n° 265171 (MP); European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 778322 (HP); Science and Engineering Research Board, New Delhi, Govt. of India (Grant No. YSS/2015/000479 dated 12 January 2016) (PS); The Chilean Government research grants Fondecyt No. 1191816 and FONDEF No. ID19 10421 (CSE); The Deutsche Forschungsgemeinschaft (DFG) Priority Program 1374 Biodiversity Exploratories (PS); European Space Agency projects IFBN (4000114425/15/NL/FF/gp) and CCI Biomass (4000123662/18/I-NB) (DS); FunDivEUROPE, European Union Seventh Framework Programme (FP7/2007–2013) under grant agreement n◦ 265171 (MS-L); APVV 20-0168 from the Slovak Research and Development Agency (VS); Manchester Metropolitan University's Environmental Science Research Centre (GS); European Union Seventh Framework Program (FP7/2007-2013) under grant agreement n° 265171 (FS); the project "LIFE+ ForBioSensing PL Comprehensive monitoring of stand dynamics in Białowieża Forest supported with remote sensing techniques" which is co-funded by the EU Life Plus programme (contract number LIFE13 ENV/PL/000048) and The National Fund for Environmental Protection and Water Management in Poland (contract number 485/2014/WN10/OP-NM-LF/D) (KJS); Global Challenges Research Fund (QR allocation, MMU) (MJPS); Czech Science Foundation project 21-27454S (MS); the Russian Foundation for Basic Research, project # 20-05-00540 (NT); Botanical Research Fund, Coalbourn Trust, Bentham Moxon Trust, Emily Holmes scholarship (LAT); the programs of the current scientific research of the Botanical Garden of the Ural Branch of Russian Academy of Sciences (VAU); FCT - Portuguese Foundation for Science and Technology - Project UIDB/04033/2020. Inventário Florestal Nacional – ICNF (HV); Grant from Kenya Coastal Development Project (KCDP), which was funded by World Bank (CW); Grants from the Swedish National Forest Inventory, Swedish University of Agricultural Sciences (BW); ATTO project (grant No. MCTI-FINEP 1759/10 and BMBF 01LB1001A, 01LK1602F) (FW); ReVaTene/PReSeD-CI 2 is funded by the Education and Research Ministry of Côte d'Ivoire, as part of the Debt Reduction-Development Contracts (C2Ds) managed by IRD (ICZ-B); the National Research Foundation of South Africa (NRF, grant 89967) (CH). The Tropical

Plant Exploration Group 70 1-ha plots in Continental Cameroon Mountains are supported by Rufford Small Grant Foundation, UK and 4-ha in Sierra Leone are supported by the Global Challenge Research Fund through Manchester Metropolitan University, UK; the National Geographic Explorer Grant, NGS-53344R-18 (ACS); University of KwaZulu-Natal Research Office grant (MJL); Universidad Nacional Autónoma de México, Dirección General de Asuntos de Personal Académico, Grant PAPIIT IN-217620 (JM). Czech Science Foundation project 21-24186M (RT, SD). The Ugandan NBS was supported with funds from the Forest Carbon Partnership Facility (FCPF), the Austrian Development Agency (ADC) and FAO. FAO's UN-REDD Program, together with the project on "Native Forests and Community"'s Loan BIRF no. 8493-AR UNDP ARG/15/004 and the National Program for the Protection of Native Forests under UNDP funded Argentina's INBN2.

Author contributions:

Conceptualization: JL, CH

Methodology: JL, CH, JGPG, NP

Data coordination: JL, MZ, SdM, TC, GJN, PBR, FS, KvG, JGPG, NP

Writing, revision, & editing: All

Competing interests

The authors declare no competing interests.

Figure Captions

Fig. 1 | A conceptual diagram of the three-stage process employed in the study. (Stage 1) Two independent global forest biodiversity individual-based (GFBi) datasets (Phase-I and Phase-II, see *Extended Data Fig. 1* for details) were standardized into a global tree-level dataframe, and aggregated into a global species abundance matrix. Based on plot locations, we merged the abundance matrix with 47 explanatory variables (*Fig. 3*) into a standardized plot-level dataframe. (**Stage 2**) We compared three candidate models (RF: random forests, XGB: XGBoost, OLS: ordinary least squares) trained from the Phase-I plot-level dataframe, using random and spatial cross-validation based on Phase-I data, and post-sample validation based on Phase-II data (Extended Data Fig. 2). The final model was then selected and re-calibrated with both Phase-I and Phase-II data. (**Stage 3**) Using the final model, we standardized and mapped local tree species richness per hectare across the global forest range. Based on this globally continuous map, we quantified the associated latitudinal diversity gradient (LDG, *Fig. 4A*), and tested for the metabolic theory of biodiversity (MTB, Fig. 2). We further developed the global map of co-limitation (*Fig. 5A*) based on model sensitivity analysis, and quantified the contribution of key factors to local species richness patterns using variance partitioning (*Fig. 6*). Dotted boxes represent processes or models, and dashed ones represent data or results. See **Methods** for details.

Fig. 2 | Latitudinal gradients of estimated tree species richness and co-limitation of drivers. (A) The latitudinal diversity gradient (LDG) of tree species richness per hectare was first empirically derived for all 0.025° pixels within the global forest range and aggregated by latitude (see Materials and Methods, data are presented as mean values +/- SD), and then compared to LDG predicted by the metabolic theory of biodiversity (MTB) based on local mean annual temperatures. **(B)** The co-limitation illustrated here was the product of LDG and the percentage prevalence of dominant drivers by latitude (**Fig. 5**).

Fig. 3 | A total of 47 explanatory variables in five categories (bioclimatic, vegetation and survey, topographic, anthropogenic, and soil) were used in random forests models to predict local tree species richness and quantify LDG. According to standardized variable importance values (horizontal bar plots to the left), bioclimatic variables contributed the most to LDG, followed by vegetation and survey, topographic, anthropogenic, and soil variables. The correlogram to the right illustrates correlations between any two variables by the color (color ramp represents the correlation coefficient) and size of a disk. The partial dependence plots to the left (next to the variable names, see **Extended Data Fig. 3** for details) show the effect of each predictor variable on the species richness, while all the other predictors remained constant at their sample mean. See **Extended Data Table 1** for a detailed description of the explanatory variables.

Fig. 4 | Estimated tree species richness per hectare in forested areas worldwide. (A) Tree species richness per hectare were first derived for the ca. 1.3 million GFBi plots across the world, and then imputed to the global forest extent. Curves (Top left) represent the observed latitudinal diversity gradient (LDG, black) of tree species diversity in comparison with LDG (red) predicted by the metabolic theory of biodiversity (MTB) based on local mean annual temperatures (see **Fig. 2**). (**B**) Width of the 95% confidence interval (C.I.) for the estimated tree species richness per hectare. All map layers are displayed at a 0.025°×0.025° resolution with an equirectangular projection (Plate-Carrée) for better illustration of the latitudinal gradients.

Fig. 5 | Dominant drivers of tree species richness in forested areas worldwide. (A) Driver dominance was derived for each pixel from four driver categories (*i.e.*, bioclimatic, topographic, anthropogenic, and soil), as well as co-limitation which represents a lack of clear dominance among the four foregoing categories. The pixel-level drivers were then aggregated by 0.5° latitudinal bins to show the percentage prevalence of dominant drivers by latitude (Top left). **(B)** The violin charts show the kernel probability density of tree species richness per hectare for different drivers. Inside boxes indicate the median (line in the center) and interquartile range (bounds of boxes). The numbers on top of the violin charts indicate the percentage of forested pixels globally that corresponds to each driver category. Red line represents the mean and 95% confidence interval of tree species richness per hectare (81.1 ± 0.1) for all the 0.025°×0.025° pixels of co-limitation. The vertical axis is on a logarithmic-10 scale for better illustration.

Fig. 6 | Patterns and variance of local tree species richness per hectare by continent. The collage of maps shows the zoomed-in view of the distribution of predicted local tree species richness per hectare (**Fig. 4A**) by continent. Circular Venn diagrams (with the legend in the center) show, for each continent, the spatial variance in observed tree species richness partitioned as follows: [*a*] (mean=14.3%) represents the fraction of variance uniquely explained by environmental factors (*i.e.* bioclimatic, topographic, anthropogenic, and soil variables), after latitudinal effects had been accounted for. [*b*] (mean=68.2%) stands for the fraction of variance jointly explained by environmental factors and latitudinal effects. [*c*] (mean=0.3%) represents the fraction of variance explained by latitudinal effects after removing environmental effects. [*d*] (mean=17.2%) represents the fraction of unexplained variance in tree species richness. The fractions were based on contrasting the amount of local richness variations in sample data from ~1.3 million plots explained by the R^2 statistics from the continental-scale random forest models with the full set of factors versus those with targeted factors removed.

References

1 Sutherland, W. J. *et al.* Identification of 100 fundamental ecological questions. *Journal of Ecology* **101**, 58-67, (2013).

2 Gaston, K. J. Global patterns in biodiversity. *Nature* **405**, 220-227, (2000).

3 Pickering, C. *The Geographical Distribution of Animals and Plants*. (Little, Brown & Co., 1854).

4 Humboldt, A. v. & Bonpland, A. Essai sur la géographie des plantes accompagné d'un tableau physique des régions équinoxiales. *Paris: Levrault, Schoell & Co*, (1805).

5 Hillebrand, H. On the generality of the latitudinal diversity gradient. *The American Naturalist* **163**, 192-211, (2004).

6 Crame, J. A. Taxonomic diversity gradients through geological time. *Diversity and Distributions* **7**, 175-189, (2001).

7 Gough, L. & Field, R. Latitudinal diversity gradients. *eLS*, (2007).

8 Willig, M. R., Kaufman, D. M. & Stevens, R. D. Latitudinal gradients of biodiversity: pattern, process, scale, and synthesis. *Annual Review of Ecology, Evolution, and Systematics* **34**, 273-309, (2003).

9 Pontarp, M. *et al.* The latitudinal diversity gradient: novel understanding through mechanistic eco-evolutionary models. *Trends in Ecology & Evolution* **34**, 211-223, (2019).

10 Vellend, M. *The theory of ecological communities (MPB-57)*. Vol. 75 (Princeton University Press, 2016).

11 Brown, J. H. Why are there so many species in the tropics? *Journal of Biogeography* **41**, 8-22, (2014).

12 Currie, D. J. *et al.* Predictions and tests of climate-based hypotheses of broad-scale variation in taxonomic richness. *Ecology Letters* **7**, 1121-1134, (2004).

13 Baldeck, C. A. *et al.* Soil resources and topography shape local tree community structure in tropical forests. *Proceedings of the Royal Society B: Biological Sciences* **280**, 20122532, (2013).

14 Qian, H. & Ricklefs, R. E. Large-scale processes and the Asian bias in species diversity of temperate plants. *Nature* **407**, 180-182, (2000).

15 Stein, A., Gerstner, K. & Kreft, H. Environmental heterogeneity as a universal driver of species richness across taxa, biomes and spatial scales. *Ecology Letters* **17**, 866-880, (2014).

16 Sponsel, L. E. Human impact on biodiversity, overview. *Encyclopedia of Biodiversity*, 137-152, (2013).

17 Sullivan, M. J. *et al.* Diversity and carbon storage across the tropical forest biome. *Scientific reports* **7**, 1-12, (2017).

18 Cazzolla Gatti, R. *et al.* The number of tree species on Earth. *Proceedings of the National Academy of Sciences* **119**, e2115329119, (2022).

19 Allen, A. P., Brown, J. H. & Gillooly, J. F. Global biodiversity, biochemical kinetics, and the energetic-equivalence rule. *Science* **297**, 1545-1548, (2002).

20 Stegen, J. C., Enquist, B. J. & Ferriere, R. Advancing the metabolic theory of biodiversity. *Ecology Letters* **12**, 1001-1015, (2009).

21 Wright, D. H. Species-energy theory: an extension of species-area theory. *Oikos* **41**, 496-506, (1983).

22 Hawkins, B. A. *et al.* Energy, water, and broad-scale geographic patterns of species richness. *Ecology* **84**, 3105-3117, (2003).

23 Staal, A., Dekker, S. C., Xu, C. & van Nes, E. H. Bistability, spatial interaction, and the distribution of tropical forests and savannas. *Ecosystems* **19**, 1080-1091, (2016).

24 de L. Dantas, V., Batalha, M. A. & Pausas, J. G. Fire drives functional thresholds on the savanna–forest transition. *Ecology* **94**, 2454-2463, (2013).

25 Bodart, C. *et al.* Continental estimates of forest cover and forest cover changes in the dry ecosystems of Africa between 1990 and 2000. *Journal of Biogeography* **40**, 1036-1047, (2013).

26 Hubau, W. *et al.* The persistence of carbon in the African forest understory. *Nature Plants* **5**, 133-140, (2019).

27 Šímová, I. & Storch, D. The enigma of terrestrial primary productivity: measurements, models, scales and the diversity–productivity relationship. *Ecography* **40**, 239-252, (2017).

28 Kreft, H. & Jetz, W. Global patterns and determinants of vascular plant diversity. *Proceedings of the National Academy of Sciences* **104**, 5925-5930, (2007).

29 Clark, D. A. *et al.* Net primary production in tropical forests: an evaluation and synthesis of existing field data. *Ecological applications* **11**, 371-384, (2001).

30 Pärtel, M., Laanisto, L. & Zobel, M. Contrasting plant productivity–diversity relationships across latitude: the role of evolutionary history. *Ecology* **88**, 1091-1097, (2007).

31 Liang, J. *et al.* Positive biodiversity-productivity relationship predominant in global forests. *Science* **354**, aaf8957, (2016).

32 Hammond, W. M. *et al.* Global field observations of tree die-off reveal hotter-drought fingerprint for Earth's forests. *Nature Communications* **13**, 1761, (2022).

33 Martens, C. *et al.* Large uncertainties in future biome changes in Africa call for flexible climate adaptation strategies. *Global Change Biology* **27**, 340-358, (2021).

34 Chapin, F. S., Bloom, A. J., Field, C. B. & Waring, R. H. Plant Responses to Multiple Environmental Factors. *BioScience* **37**, 49-57, (1987).

35 Tilman, D. *Resource Competition and Community Structure.(MPB-17), Volume 17*. (Princeton university press, 1982).

36 Harpole, W. S. *et al.* Nutrient co‐limitation of primary producer communities. *Ecology letters* **14**, 852-862, (2011).

37 Colwell, R. K., Rahbek, C. & Gotelli, N. J. The mid-domain effect and species richness patterns: what have we learned so far? *The American Naturalist* **163**, E1-E23, (2004).

38 Feng, G. *et al.* Species and phylogenetic endemism in angiosperm trees across the Northern Hemisphere are jointly shaped by modern climate and glacial–interglacial climate change. *Global Ecology and Biogeography* **28**, 1393-1402, (2019).

39 Fraser, C. I., Nikula, R., Ruzzante, D. E. & Waters, J. M. Poleward bound: biological impacts of Southern Hemisphere glaciation. *Trends in Ecology & Evolution* **27**, 462-471, (2012).

40 Algar, A. C., Kerr, J. T. & Currie, D. J. A test of metabolic theory as the mechanism underlying broad‐scale species‐richness gradients. *Global Ecology and Biogeography* **16**, 170-178, (2007).

41 Wilson, E. O. A Global Biodiversity Map. *Science* **289**, 2279-2279, (2000).

42 Castagneyrol, B. & Jactel, H. Unraveling plant-animal diversity relationships: a meta-regression analysis. *Ecology* **93**, 2115-2124, (2012).

43 Steidinger, B. S. *et al.* Climatic controls of decomposition drive the global biogeography of forest-tree symbioses. *Nature* **569**, 404-408, (2019).

44 Bongers, F. J. *et al.* Functional diversity effects on productivity increase with age in a forest biodiversity experiment. *Nature Ecology & Evolution* **5**, 1594-1603, (2021).

45 UN Statistical Commission. System of Environmental-Economic Accounting—Ecosystem Accounting: Final Draft. 350 (United Nations, 2021).

46 The World Bank. Green Bond Impact Report. (Capital Markets Department, the World Bank Treasury, Washington, DC, 2019).

47 Crowther, T. *et al.* Mapping tree density at a global scale. *Nature* **525**, 201-205, (2015).

48 Chamberlain, J. L., Prisley, S. & McGuffin, M. Understanding the relationships between American ginseng harvest and hardwood forests inventory and timber harvest to improve co-management of the forests of eastern United States. *Journal of sustainable forestry* **32**, 605-624, (2013).

49 R Core Team. R: A language and environment for statistical computing. (R Foundation for Statistical Computing, Vienna, Austria, 2020).

50 Borda-de-Água, L., Hubbell, S. P. & McAllister, M. Species-area curves, diversity indices, and species abundance distributions: a multifractal analysis. *American Naturalist* **159**, 138-155, (2002).

51 Connor, E. F. & McCoy, E. D. The statistics and biology of the species-area relationship. *American Naturalist* **113**, 791-833, (1979).

52 He, F. & Legendre, P. On species-area relations. *American Naturalist* **148**, 719-737, (1996).

53 White, E. P., Ernest, S. K. M., Kerkhoff, A. J. & Enquist, B. J. Relationships between body size and abundance in ecology. *Trends in Ecology & Evolution* **22**, 323-330, (2007).

54 Gotelli, N. J. & Colwell, R. K. Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. *Ecology Letters* **4**, 379-391, (2001).

55 Chirici, G., Winter, S. & McRoberts, R. E. *National forest inventories: Contributions to forest biodiversity assessments*. Vol. 20 (Springer Science & Business Media, 2011).

56 Tomppo, E. *et al. National forest inventories: Pathways for Common Reporting*. (Springer, 2010).

57 Winter, S., Böck, A. & McRoberts, R. E. Estimating tree species diversity across geographic scales. *European journal of forest research* **131**, 441-451, (2012).

58 McRoberts, R. E., Winter, S., Chirici, G. & LaPoint, E. Assessing forest naturalness. *Forest Science* **58**, 294-309, (2012).

59 ESRI. Release 10.3 of Desktop, ESRI ArcGIS. (Environmental Systems Research Institute, Redlands, CA, 2014).

60 Breiman, L. Random forests. *Machine Learning* **45**, 5-32, (2001).

61 Strobl, C., Boulesteix, A.-L., Kneib, T., Augustin, T. & Zeileis, A. Conditional variable importance for random forests. *BMC Bioinformatics* **9**, 307, (2008).

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

63 Hyndman, R. J. & Koehler, A. B. Another look at measures of forecast accuracy. *International Journal of Forecasting* **22**, 679-688, (2006).

64 Legendre, P. Spatial autocorrelation: trouble or new paradigm? *Ecology* **74**, 1659-1673, (1993).

65 Husch, B., Beers, T. W. & Kershaw Jr, J. A. *Forest mensuration*. 4th edn, (John Wiley & Sons, 2003).

66 Farrar, D. E. & Glauber, R. R. Multicollinearity in regression analysis: the problem revisited. *The Review of Economics and Statistics*, 92-107, (1967).

67 Venables, W. N. & Ripley, B. D. *Modern applied statistics with S-PLUS*. (Springer Science & Business Media, 2013).

68 Chen, T. & Guestrin, C. in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 785-794 (ACM).

69 Friedman, J. H. Greedy function approximation: a gradient boosting machine. *Annals of Statistics*, 1189-1232, (2001).

70 Hansen, M. C. *et al.* High-resolution global maps of 21st-century forest cover change. *Science* **342**, 850-853, (2013).

71 FAO. Global Forest Resources Assessment 2015 - How are the world's forests changing? , (Food and Agriculture Organization of the United Nations, Rome, Italy, 2015).

72 Bastin, J.-F. *et al.* The extent of forest in dryland biomes. *Science* **356**, 635-638, (2017).

73 Peres-Neto, P. R., Legendre, P., Dray, S. & Borcard, D. Variation partitioning of species data matrices: estimation and comparison of fractions. *Ecology* **87**, 2614-2625, (2006).

74 Peres-Neto, P. R. & Legendre, P. Estimating and controlling for spatial structure in the study of ecological communities. *Global Ecology and Biogeography* **19**, 174-184, (2010).

75 Saltelli, A. *et al. Global sensitivity analysis: the primer*. (John Wiley & Sons, 2008).

76 Liang, J. & Gamarra, J. G. P. The importance of sharing global forest data in a world of crises. *Scientific Data* **7**, 424, (2020).

77 FAO. Towards open and transparent forest data for climate action - Experiences and lessons learned. (United Nations' Food and Agriculture Organization, Rome, Italy, 2022).