



OPEN

Unraveling Amazon tree community assembly using Maximum Information Entropy: a quantitative analysis of tropical forest ecology

Edwin Pos^{1,2,✉}, Luiz de Souza Coelho³, Diogenes de Andrade Lima Filho³, Rafael P. Salomão^{4,5}, Iêda Leão Amaral³, Francisca Dionízia de Almeida Matos³, Carolina V. Castilho⁶, Oliver L. Phillips⁷, Juan Ernesto Guevara^{8,9}, Marcelo de Jesus Veiga Carim¹⁰, Dairon Cárdenas López¹¹, William E. Magnusson¹², Florian Wittmann^{13,14}, Mariana Victória Irumé³, Maria Pires Martins³, Daniel Sabatier¹⁵, José Renan da Silva Guimarães¹⁰, Jean-François Molino¹⁵, Olaf S. Bánki², Maria Teresa Fernandez Piedade¹⁶, Nigel C. A. Pitman¹⁷, Abel Monteagudo Mendoza¹⁸, José Ferreira Ramos³, Joseph E. Hawes¹⁹, Everton José Almeida²⁰, Luciane Ferreira Barbosa²⁰, Larissa Cavalheiro²⁰, Márcia Cléia Vilela dos Santos²⁰, Bruno Garcia Luize²¹, Evelyn Márcia Moraes de Leão Novo²², Percy Núñez Vargas²³, Thiago Sanna Freire Silva²⁴, Eduardo Martins Venticinque²⁵, Angelo Gilberto Manzatto²⁶, Neidiane Farias Costa Reis²⁷, John Terborgh^{28,29}, Katia Regina Casula²⁷, Euridice N. Honorio Coronado^{7,30}, Juan Carlos Montero^{3,31}, Beatriz S. Marimon³², Ben Hur Marimon-Junior³², Ted R. Feldpausch^{7,33}, Alvaro Duque³⁴, Chris Baraloto³⁵, Nicolás Castaño Arboleda¹¹, Julien Engel^{15,35}, Pascal Petronelli³⁶, Charles Eugene Zartman³, Timothy J. Killeen³⁷, Rodolfo Vasquez¹⁸, Bonifacio Mostacedo³⁸, Rafael L. Assis³⁹, Jochen Schöngart¹⁶, Hernán Castellanos⁴⁰, Marcelo Brilhante de Medeiros⁴¹, Marcelo Fragomeni Simon⁴¹, Ana Andrade⁴², José Luís Camargo⁴², Layon O. Demarchi¹⁶, William F. Laurance²⁹, Susan G. W. Laurance²⁹, Emanuelle de Sousa Farias^{43,44}, Maria Aparecida Lopes⁴⁵, José Leonardo Lima Magalhães^{46,47}, Henrique Eduardo Mendonça Nascimento³, Helder Lima de Queiroz⁴⁸, Gerardo A. C. Aymard⁴⁹, Roel Brienen⁷, Juan David Cardenas Revilla³, Flávia R. C. Costa³, Adriano Quaresma¹⁶, Ima Célia Guimarães Vieira⁵, Bruno Barçante Ladvoat Cintra⁵⁰, Pablo R. Stevenson⁵¹, Yuri Oliveira Feitosa⁵², Joost F. Duivenvoorden⁵³, Hugo F. Mogollón⁵⁴, Leandro Valle Ferreira⁵, James A. Comiskey^{55,56}, Freddie Draper^{35,57}, José Julio de Toledo⁵⁸, Gabriel Damasco⁵⁹, Nállarett Dávila⁶⁰, Roosevelt García-Villacorta^{61,62}, Aline Lopes^{16,63}, Alberto Vicentini¹², Janaina Costa Noronha⁶⁴, Flávia Rodrigues Barbosa⁶⁴, Rainiellen de Sá Carpanedo⁶⁴, Thaise Emilio^{12,65}, Carolina Levis^{66,67}, Domingos de Jesus Rodrigues⁶⁴, Juliana Schietti³, Priscila Souza³, Alfonso Alonso⁵⁶, Francisco Dallmeier⁵⁶, Vitor H. F. Gomes^{68,69}, Jon Lloyd⁷⁰, David Neill⁷¹, Daniel Praia Portela de Aguiar¹⁶, Alejandro Araujo-Murakami⁷², Luzmila Arroyo⁷², Fernanda Antunes Carvalho^{12,73}, Fernanda Coelho de Souza^{7,12}, Dário Dantas do Amaral⁵, Kenneth J. Feeley^{74,75}, Rogerio Gribel³, Marcelo Petratti Pansonato^{3,76}, Jos Barlow⁷⁷, Erika Berenguer^{77,78}, Joice Ferreira⁴⁷, Paul V. A. Fine⁵⁹, Marcelino Carneiro Guedes⁷⁹, Eliana M. Jimenez⁸⁰, Juan Carlos Licona³¹, Maria Cristina Peñuela Mora⁸¹, Carlos A. Peres⁸², Boris Eduardo Villa Zegarra⁸³, Carlos Cerón⁸⁴, Terry W. Henkel⁸⁵, Paul Maas², Marcos Silveira⁸⁶, Juliana Stropp⁸⁷, Raquel Thomas-Caesar⁸⁸, Tim R. Baker⁷, Doug Daly⁸⁹, Kyle G. Dexter^{90,91}, John Ethan Householder¹³, Isau Huamantupa-Chuquimaco²³, Toby Pennington^{33,91}, Marcos Ríos Paredes⁹², Alfredo Fuentes^{93,94}, José Luis Marcelo Pena⁹⁵, Miles R. Silman⁹⁶, J. Sebastián Tello⁹⁴, Jerome Chave⁹⁷, Fernando Cornejo Valverde⁹⁸,

Anthony Di Fiore⁹⁹, Renato Richard Hilário⁵⁸, Juan Fernando Phillips¹⁰⁰, Gonzalo Rivas-Torres^{101,102}, Tinde R. van Andel^{2,103}, Patricio von Hildebrand¹⁰⁴, Edelcilio Marques Barbosa³, Luiz Carlos de Matos Bonates³, Hilda Paulette Dávila Doza⁹², Émile Fonty^{15,105}, Ricardo Zárate Gómez¹⁰⁶, Therany Gonzales¹⁰⁷, George Pepe Gallardo Gonzales⁹², Jean-Louis Guillaumet^{108,145}, Bruce Hoffman¹⁰⁹, André Braga Junqueira¹¹⁰, Yadvinder Malhi¹¹¹, Ires Paula de Andrade Miranda³, Linder Felipe Mozombite Pinto⁹², Adriana Prieto¹¹², Agustín Rudas¹¹², Ademir R. Ruschel⁴⁷, Natalino Silva¹¹³, César I. A. Vela¹¹⁴, Vincent Antoine Vos¹¹⁵, Eglé L. Zent¹¹⁶, Stanford Zent¹¹⁶, Bianca Weiss Albuquerque¹⁶, Angela Cano^{51,117}, Diego F. Correa^{51,118}, Janaina Barbosa Pedrosa Costa⁷⁹, Bernardo Monteiro Flores¹¹⁹, Milena Holmgren¹²⁰, Marcelo Trindade Nascimento¹²¹, Alexandre A. Oliveira⁷⁶, Hirma Ramirez-Angulo¹²², Maira Rocha¹⁶, Veridiana Vizoni Scudeller¹²³, Rodrigo Sierra¹²⁴, Milton Tirado¹²⁴, Maria Natalia Umaña¹²⁵, Geertje van der Heijden¹²⁶, Emilio Vilanova Torre^{122,127}, Corine Vriesendorp¹⁷, Ophelia Wang¹²⁸, Kenneth R. Young¹²⁹, Manuel Augusto Ahuite Reategui¹³⁰, Cláudia Baider^{76,131}, Henrik Balslev¹³², Sasha Cárdenas⁵¹, Luisa Fernanda Casas⁵¹, William Farfan-Rios^{23,94,133}, Cid Ferreira³, Reynaldo Linares-Palomino⁵⁶, Casimiro Mendoza^{134,135}, Italo Mesones⁵⁹, Armando Torres-Lezama¹²², Ligia Estela Urrego Giraldo³⁴, Daniel Villarroya⁷², Roderick Zagt¹³⁶, Miguel N. Alexiades¹³⁷, Karina Garcia-Cabrera⁹⁶, Lionel Hernandez⁴⁰, William Milliken⁶⁵, Walter Palacios Cuenca¹³⁸, Susamar Pansini²⁷, Daniela Pauletto¹³⁹, Freddy Ramirez Arevalo¹⁴⁰, Adeilza Felipe Sampaio²⁷, Elvis H. Valderrama Sandoval^{140,141}, Luis Valenzuela Gamarra¹⁸, Gerhard Boenisch¹⁴², Jens Kattge¹⁴³, Nathan Kraft¹⁴⁴, Aurora Levesley⁷, Karina Melgaço⁷, Georgia Pickavance⁷, Lourens Poorter⁶⁷ & Hans ter Steege²

In a time of rapid global change, the question of what determines patterns in species abundance distribution remains a priority for understanding the complex dynamics of ecosystems. The constrained maximization of information entropy provides a framework for the understanding of such complex systems dynamics by a quantitative analysis of important constraints via predictions using least biased probability distributions. We apply it to over two thousand hectares of Amazonian tree inventories across seven forest types and thirteen functional traits, representing major global axes of plant strategies. Results show that constraints formed by regional relative abundances of genera explain eight times more of local relative abundances than constraints based on directional selection for specific functional traits, although the latter does show clear signals of environmental dependency. These results provide a quantitative insight by inference from large-scale data using cross-disciplinary methods, furthering our understanding of ecological dynamics.

¹Quantitative Biodiversity Dynamics, Ecology and Biodiversity, Utrecht University Botanic Gardens, Utrecht University, Padualaan 8, Utrecht 3584 CH, The Netherlands. ²Naturalis Biodiversity Center, PO Box 9517, Leiden 2300 RA, The Netherlands. ³Coordenação de Biodiversidade, Instituto Nacional de Pesquisas da Amazônia - INPA, Av. André Araújo, 2936, Petrópolis, Manaus, AM 69067-375, Brazil. ⁴Programa Professor Visitante Nacional Sênior Na Amazônia - CAPES, Universidade Federal Rural da Amazônia, Av. Perimetral, s/n, Belém, PA, Brazil. ⁵Coordenação de Botânica, Museu Paraense Emílio Goeldi, Av. Magalhães Barata 376, C.P. 399, Belém, PA 66040-170, Brazil. ⁶EMBRAPA - Centro de Pesquisa Agroflorestal de Roraima, BR 174, km 8 - Distrito Industrial, Boa Vista, RR 69301-970, Brazil. ⁷School of Geography, University of Leeds, Woodhouse Lane, Leeds LS2 9JT, UK. ⁸Grupo de Investigación en Biodiversidad, Medio Ambiente y Salud-BIOMAS, Universidad de Las Américas, Campus Queri, Quito, Ecuador. ⁹Keller Science Action Center, The Field Museum, 1400 S. Lake Shore Drive, Chicago, IL 60605-2496, USA. ¹⁰Departamento de Botânica, Instituto de Pesquisas Científicas e Tecnológicas do Amapá - IEPA, Rodovia JK, Km 10, Campus Do IEPA da Fazendinha, Amapá 68901-025, Brazil. ¹¹Herbario Amazônico Colombiano, Instituto SINCHI, Calle 20 No 5-44, Bogotá, DC, Colombia. ¹²Coordenação de Pesquisas em Ecologia, Instituto Nacional de Pesquisas da Amazônia - INPA, Av. André Araújo, 2936, Petrópolis, Manaus, AM 69067-375, Brazil. ¹³Department of Wetland Ecology, Institute of Geography and Geoecology, Karlsruhe Institute of Technology - KIT, Josefstr.1, 76437 Rastatt, Germany. ¹⁴Biogeochemistry, Max Planck Institute for Chemistry, Hahn-Meitner Weg 1, 55128 Mainz, Germany. ¹⁵AMAP, IRD, Cirad, CNRS, INRA, Université de Montpellier, 34398 Montpellier, France. ¹⁶Coordenação de Dinâmica Ambiental, Instituto Nacional de Pesquisas da Amazônia - INPA, Av. André Araújo, 2936, Petrópolis, Manaus, AM 69067-375, Brazil. ¹⁷Science and Education, The Field Museum, 1400 S. Lake Shore Drive, Chicago, IL 60605-2496, USA. ¹⁸Jardín Botánico de Missouri, Oxapampa, Pasco, Peru. ¹⁹Applied Ecology Research Group, School of Life Sciences, Anglia Ruskin University, East Road, Cambridge CB1 1PT, UK. ²⁰ICNHS, Universidade Federal de Mato Grosso, Av. Alexandre Ferronato, 1200, Sinop, MT 78557-267, Brazil. ²¹Departamento de Ecologia, Universidade Estadual Paulista - UNESP - Instituto de Biociências - IB, Av. 24 A, 1515, Bela Vista, Rio Claro, SP 13506-900, Brazil. ²²Divisão de Sensoriamento Remoto - DSR, Instituto Nacional de Pesquisas Espaciais - INPE, Av. Dos Astronautas, 1758, Jardim da Granja, São José Dos

Campos, SP 12227-010, Brazil. ²³Herbario Vargas, Universidad Nacional de San Antonio Abad del Cusco, Avenida de La Cultura, Nro 733, Cusco, Cuzco, Peru. ²⁴Biological and Environmental Sciences, University of Stirling, Stirling FK9 4LA, UK. ²⁵Centro de Biociências, Departamento de Ecologia, Universidade Federal do Rio Grande do Norte, Av. Senador Salgado Filho, 3000, Natal, RN 59072-970, Brazil. ²⁶Departamento de Biologia, Universidade Federal de Rondônia, Rodovia BR 364 s/n Km 9, 5 - Sentido Acre, Unir, Porto Velho, RO 76.824-027, Brazil. ²⁷Programa de Pós- Graduação em Biodiversidade e Biotecnologia PPG- Bionorte, Universidade Federal de Rondônia, Campus Porto Velho Km 9, 5 Bairro Rural, Porto Velho, RO 76.824-027, Brazil. ²⁸Department of Biology and Florida Museum of Natural History, University of Florida, Gainesville, FL 32611, USA. ²⁹Centre for Tropical Environmental and Sustainability Science and College of Science and Engineering, James Cook University, Cairns, QLD 4870, Australia. ³⁰Instituto de Investigaciones de la Amazonía Peruana (IIAP), Av. A. Quiñones Km 2,5, Iquitos, Loreto 784, Peru. ³¹Instituto Boliviano de Investigación Forestal, Av. 6 de Agosto #28, Km. 14, Doble via La Guardia, 6204 Santa Cruz, Santa Cruz, Casilla, Bolivia. ³²Programa de Pós-Graduação em Ecologia e Conservação, Universidade do Estado de Mato Grosso, Nova Xavantina, MT, Brazil. ³³Geography, College of Life and Environmental Sciences, University of Exeter, Rennes Drive, Exeter EX4 4RJ, UK. ³⁴Departamento de Ciencias Forestales, Universidad Nacional de Colombia, Calle 64 X Cra 65, 1027 Medellín, Antioquia, Colombia. ³⁵International Center for Tropical Botany (ICTB) Department of Biological Sciences, Florida International University, 11200 SW 8Th Street, OE 243, Miami, FL 33199, USA. ³⁶Cirad UMR Ecofog, AgrosParisTech, CNRS, INRA, Univ Guyane, Campus Agronomique, 97379 Kourou Cedex, France. ³⁷Agteca-Amazonica, Santa Cruz, Bolivia. ³⁸Facultad de Ciencias Agrícolas, Universidad Autónoma Gabriel René Moreno, Santa Cruz, Santa Cruz, Bolivia. ³⁹Natural History Museum, University of Oslo, Postboks 1172, 0318 Oslo, Norway. ⁴⁰Centro de Investigaciones Ecológicas de Guayana, Universidad Nacional Experimental de Guayana, Calle Chile, Urbaniz Chilemex, Puerto Ordaz, Bolivar, Venezuela. ⁴¹Prédio da Botânica e Ecologia, Embrapa Recursos Genéticos e Biotecnologia, Parque Estação Biológica, Av. W5 Norte, Brasília, DF 70770-917, Brazil. ⁴²Projeto Dinâmica Biológica de Fragmentos Florestais, Instituto Nacional de Pesquisas da Amazônia - INPA, Av. André Araújo 2936, Petrópolis, Manaus, AM 69067-375, Brazil. ⁴³Laboratório de Ecologia de Doenças Transmissíveis da Amazônia (EDTA), Instituto Leônidas e Maria Deane, Fiocruz, Rua Terezina, 476, Adrianópolis, Manaus, AM 69060-001, Brazil. ⁴⁴Programa de Pós-Graduação em Biodiversidade e Saúde, Instituto Oswaldo Cruz - IOC/FIOCRUZ, Pav. Arthur Neiva – Térreo, Av. Brasil, 4365 – Manguinhos, Rio de Janeiro, RJ 21040-360, Brazil. ⁴⁵Instituto de Ciências Biológicas, Universidade Federal do Pará, Av. Augusto Corrêa 01, Belém, PA 66075-110, Brazil. ⁴⁶Programa de Pós-Graduação em Ecologia, Universidade Federal do Pará, Av. Augusto Corrêa 01, Belém, PA 66075-110, Brazil. ⁴⁷Embrapa Amazônia Oriental, Trav. Dr. Enéas Pinheiro S/nº, Belém, PA 66095-100, Brazil. ⁴⁸Diretoria Técnico-Científica, Instituto de Desenvolvimento Sustentável Mamirauá, Estrada do Bexiga, 2584, Tefé, AM 69470-000, Brazil. ⁴⁹Programa de Ciencias del Agro y el Mar, Herbario Universitario (PORT), UNELLEZ-Guanare, Guanare, Portuguesa 3350, Venezuela. ⁵⁰Instituto de Biociências – Department of Botânica, Universidade de Sao Paulo - USP, Rua do Matão 277, Cidade Universitária, São Paulo, SP 05508-090, Brazil. ⁵¹Laboratorio de Ecología de Bosques Tropicales y Primatología, Universidad de los Andes, Carrera 1 # 18a-10, 111711 Bogotá, DC, Colombia. ⁵²Programa de Pós-Graduação Em Biologia (Botânica), Instituto Nacional de Pesquisas da Amazônia - INPA, Av. André Araújo, 2936, Petrópolis, Manaus, AM 69067-375, Brazil. ⁵³Institute of Biodiversity and Ecosystem Dynamics, University of Amsterdam, Sciencepark 904, Amsterdam 1098 XH, The Netherlands. ⁵⁴Endangered Species Coalition, 8530 Geren Rd., Silver Spring, MD 20901, USA. ⁵⁵Inventory and Monitoring Program, National Park Service, 120 Chatham Lane, Fredericksburg, VA 22405, USA. ⁵⁶Center for Conservation and Sustainability, Smithsonian Conservation Biology Institute, 1100 Jefferson Dr. SW, Suite 3123, Washington, DC 20560-0705, USA. ⁵⁷Department of Global Ecology, Carnegie Institution for Science, 260 Panama St., Stanford, CA 94305, USA. ⁵⁸Universidade Federal do Amapá, Ciências Ambientais, Rod. Juscelino Kubitschek km2, Macapá, AP 68902-280, Brazil. ⁵⁹Department of Integrative Biology, University of California, Berkeley, CA 94720-3140, USA. ⁶⁰Biologia Vegetal, Universidade Estadual de Campinas, Caixa Postal 6109, Campinas, SP 13.083-970, Brazil. ⁶¹Department of Ecology and Evolutionary Biology, Cornell University, Corson Hall, 215 Tower Road, Ithaca, NY 14850, USA. ⁶²Peruvian Center for Biodiversity and Conservation (PCBC), Iquitos, Peru. ⁶³Department of Ecology, University of Brasilia, Brasilia, DF 70904-970, Brazil. ⁶⁴ICNHS, Federal University of Mato Grosso, Av. Alexandre Ferronato 1200, Setor Industrial, Sinop, MT 78.557-267, Brazil. ⁶⁵Natural Capital and Plant Health, Royal Botanic Gardens, Kew, Richmond TW9 3AB, Surrey, UK. ⁶⁶Programa de Pós-Graduação em Ecologia, Instituto Nacional de Pesquisas da Amazônia - INPA, Av. André Araújo, 2936, Petrópolis, Manaus, AM 69067-375, Brazil. ⁶⁷Forest Ecology and Forest Management Group, Wageningen University and Research, Droevendaalsesteeg 3, P.O. Box 47, Wageningen 6700 AA, The Netherlands. ⁶⁸Escola de Negócios Tecnologia e Inovação, Centro Universitário do Pará, Belém, PA, Brazil. ⁶⁹Universidade Federal do Pará, Rua Augusto Corrêa 01, Belém, PA 66075-110, Brazil. ⁷⁰Faculty of Natural Sciences, Department of Life Sciences, Imperial College London, South Kensington Campus, Silwood Park London SW7 2AZ, UK. ⁷¹Ecosistemas, Biodiversidad y Conservación de Especies, Universidad Estatal Amazónica, Km. 2 1/2 Vía a Tena (Paso Lateral), Puyo, Pastaza, Ecuador. ⁷²Museo de Historia Natural Noel Kempff Mercado, Universidad Autónoma Gabriel Rene Moreno, Avenida Irala 565 Casilla Post Al 2489, Santa Cruz, Santa Cruz, Bolivia. ⁷³Departamento de Genética, Ecologia e Evolução, Universidade Federal de Minas Gerais, Instituto de Ciências Biológicas, Av. Antônio Carlos, 6627 Pampulha, Belo Horizonte, MG 31270-901, Brazil. ⁷⁴Department of Biology, University of Miami, Coral Gables, FL 33146, USA. ⁷⁵Fairchild Tropical Botanic Garden, Coral Gables, FL 33156, USA. ⁷⁶Instituto de Biociências - Dept. Ecologia, Universidade de Sao Paulo - USP, Rua do Matão, Trav. 14, No. 321, Cidade Universitária, São Paulo, SP 05508-090, Brazil. ⁷⁷Lancaster Environment Centre, Lancaster University, Lancaster LA1 4YQ, Lancashire, UK. ⁷⁸Environmental Change Institute, University of Oxford, Oxford OX1 3QY, Oxfordshire, UK. ⁷⁹Empresa Brasileira de Pesquisa Agropecuária, Embrapa Amapá, Rod. Juscelino Kubitschek Km 5, Macapá, Amapá 68903-419, Brazil. ⁸⁰Grupo de Ecología y Conservación de Fauna y Flora Silvestre, Instituto Amazónico de Investigaciones Imani, Universidad Nacional de Colombia Sede Amazonia, Leticia, Amazonas, Colombia. ⁸¹Universidad Regional Amazónica IKIAM, Km 7 Vía Muyuna, Tena, Napo,

Ecuador. ⁸²School of Environmental Sciences, University of East Anglia, Norwich NR4 7TJ, UK. ⁸³Dirección de Evaluación Forestal y de Fauna Silvestre, Av. Javier Praod Oeste 693, Magdalena del Mar, Peru. ⁸⁴Escuela de Biología Herbario Alfredo Paredes, Universidad Central, Ap. Postal 17.01.2177, Quito, Pichincha, Ecuador. ⁸⁵Department of Biological Sciences, Humboldt State University, 1 Harpst Street, Arcata, CA 95521, USA. ⁸⁶Museu Universitário / Centro de Ciências Biológicas e da Natureza / Laboratório de Botânica e Ecologia Vegetal, Universidade Federal do Acre, Rio Branco, AC 69915-559, Brazil. ⁸⁷Institute of Biological and Health Sciences, Federal University of Alagoas, Av. Lourival Melo Mota, s/n, Tabuleiro do Martins, Maceio, AL 57072-970, Brazil. ⁸⁸Iwokrama International Centre for Rain Forest Conservation and Development, Georgetown, Guyana. ⁸⁹New York Botanical Garden, 2900 Southern Blvd, Bronx, New York, NY 10458-5126, USA. ⁹⁰School of Geosciences, University of Edinburgh, 201 Crew Building, King's Buildings, Edinburgh EH9 3JN, UK. ⁹¹Tropical Diversity Section, Royal Botanic Garden Edinburgh, 20a Inverleith Row, Edinburgh EH3 5LR, Scotland, UK. ⁹²Servicios de Biodiversidad EIRL, Jr. Independencia 405, Iquitos, Loreto 784, Peru. ⁹³Herbario Nacional de Bolivia, Universitario UMSA, Casilla 10077 Correo Central, La Paz, La Paz, Bolivia. ⁹⁴Center for Conservation and Sustainable Development, Missouri Botanical Garden, P.O. Box 299, St. Louis, MO 63166-0299, USA. ⁹⁵Universidad Nacional de Jaén, Carretera Jaén San Ignacio Km 23, Jaén, Cajamarca 06801, Peru. ⁹⁶Biology Department and Center for Energy, Environment and Sustainability, Wake Forest University, 1834 Wake Forest Rd, Winston Salem, NC 27106, USA. ⁹⁷Laboratoire Evolution et Diversité Biologique, CNRS and Université Paul Sabatier, UMR 5174 EDB, 31000 Toulouse, France. ⁹⁸Andes to Amazon Biodiversity Program, Madre de Dios, Madre de Dios, Peru. ⁹⁹Department of Anthropology, University of Texas at Austin, SAC 5.150, 2201 Speedway Stop C3200, Austin, TX 78712, USA. ¹⁰⁰Fundación Puerto Rastrojo, Cra 10 No. 24-76 Oficina 1201, Bogotá, DC, Colombia. ¹⁰¹Colegio de Ciencias Biológicas y Ambientales-COCIBA and Galapagos Institute for the Arts and Sciences-GAIAS, Universidad San Francisco de Quito-USFQ, Quito, Pichincha, Ecuador. ¹⁰²Department of Wildlife Ecology and Conservation, University of Florida, 110 Newins-Ziegler Hall, Gainesville, FL 32611, USA. ¹⁰³Biosystematics Group, Wageningen University, Droevendaalsesteeg 1, Wageningen 6708 PB, The Netherlands. ¹⁰⁴Fundación Estación de Biología, Cra 10 No. 24-76 Oficina, 1201 Bogotá, DC, Colombia. ¹⁰⁵Direction Régionale de la Guyane, ONF, Cayenne 97300, French Guiana. ¹⁰⁶PROTERRA, Instituto de Investigaciones de la Amazonía Peruana (IIAP), Av. A. Quiñones Km 2,5, Iquitos, Loreto 784, Peru. ¹⁰⁷ACEER Foundation, Jirón Cusco N° 370, Puerto Maldonado, Madre de Dios, Peru. ¹⁰⁸Departement EV, Muséum National d'histoire Naturelle de Paris, 16 Rue Buffon, Paris 75005, France. ¹⁰⁹Amazon Conservation Team, Doekhieweg Oost #24, Paramaribo, Suriname. ¹¹⁰Institut de Ciència i Tecnologia Ambientals, Universitat Autònoma de Barcelona, 08193 Bellaterra, Barcelona, Spain. ¹¹¹Environmental Change Institute, Dyson Perrins Building, Oxford University Centre for the Environment, South Parks Road, Oxford OX1 3QY, England, UK. ¹¹²Instituto de Ciencias Naturales, Universidad Nacional de Colombia, 7945 Apartado, Bogotá, DC, Colombia. ¹¹³Instituto de Ciência Agrárias, Universidade Federal Rural da Amazônia, Av. Presidente Tancredo Neves 2501, Belém, PA 66.077-830, Brazil. ¹¹⁴Escuela Profesional de Ingeniería Forestal, Universidad Nacional de San Antonio Abad del Cusco, Jirón San Martín 451, Puerto Maldonado, Madre de Dios, Peru. ¹¹⁵Universidad Autónoma del Beni José Ballivián, Campus Universitario Final, Av. Ejercito, Riberalta, Beni, Bolivia. ¹¹⁶Laboratory of Human Ecology, Instituto Venezolano de Investigaciones Científicas - IVIC, Ado 20632, Caracas 1020A, DC, Venezuela. ¹¹⁷Cambridge University Botanic Garden, 1 Brookside., Cambridge CB2 1JE, UK. ¹¹⁸School of Agriculture and Food Sciences - ARC Centre of Excellence for Environmental Decisions CEED, The University of Queensland, St. Lucia, QLD 4072, Australia. ¹¹⁹Plant Biology Department, Rua Monteiro Lobato, University of Campinas, 255, Cidade Universitária Zeferino Vaz, Barão Geraldo, Campinas, São Paulo CEP 13083-862, Brazil. ¹²⁰Resource Ecology Group, Wageningen University and Research, Droevendaalsesteeg 3a, Lumen, Building Number 100, Wageningen, Gelderland 6708 PB, The Netherlands. ¹²¹Laboratório de Ciências Ambientais, Universidade Estadual do Norte Fluminense, Av. Alberto Lamego 2000, Campos dos, Goytacazes, RJ 28013-620, Brazil. ¹²²Instituto de Investigaciones Para el Desarrollo Forestal (INDEFOR), Universidad de los Andes, Conjunto Forestal, Mérida, Mérida 5101, Venezuela. ¹²³Departamento de Biología, Universidade Federal do Amazonas - UFAM - Instituto de Ciências Biológicas - ICB1, Av General Rodrigo Octavio 6200, Manaus, AM 69080-900, Brazil. ¹²⁴GeolS, el Día 369 y el Telégrafo, 3° Piso, Quito, Pichincha, Ecuador. ¹²⁵Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, MI 48109, USA. ¹²⁶University of Nottingham, University Park, Nottingham NG7 2RD, UK. ¹²⁷School of Environmental and Forest Sciences, University of Washington, Seattle, WA 98195-2100, USA. ¹²⁸Environmental Science and Policy, Northern Arizona University, Flagstaff, AZ 86011, USA. ¹²⁹Geography and the Environment, University of Texas at Austin, 305 E. 23Rd Street, CLA Building, Austin, TX 78712, USA. ¹³⁰Medio Ambiente, PLUSPRETOL, Iquitos, Loreto, Peru. ¹³¹The Mauritius Herbarium, Agricultural Services, Ministry of Agro-Industry and Food Security, Reduit 80835, Mauritius. ¹³²Department of Bioscience, Aarhus University, Building 1540 Ny Munkegade, 8000 Aarhus C, Aarhus, Denmark. ¹³³Living Earth Collaborative, Washington University in Saint Louis, St. Louis, MO 63130, USA. ¹³⁴Escuela de Ciencias Forestales (ESFOR), Universidad Mayor de San Simon (UMSS), Sacta, Cochabamba, Bolivia. ¹³⁵FOMABO, Manejo Forestal en Las Tierras Tropicales de Bolivia, Sacta, Cochabamba, Bolivia. ¹³⁶Tropenbos International, Lawickse Allee 11, PO Box 232, Wageningen 6700 AE, The Netherlands. ¹³⁷School of Anthropology and Conservation, University of Kent, Marlowe Building, Canterbury, Kent CT2 7NR, UK. ¹³⁸Herbario Nacional del Ecuador, Universidad Técnica del Norte, Quito, Pichincha, Ecuador. ¹³⁹Instituto de Biodiversidade e Floresta, Universidade Federal do Oeste do Pará, Rua Vera Paz, Campus Tapajós, Santarém, PA 68015-110, Brazil. ¹⁴⁰Facultad de Biología, Universidad Nacional de la Amazonia Peruana, Pevas 5Ta Cdra, Iquitos, Loreto, Peru. ¹⁴¹Department of Biology, University of Missouri, St. Louis, MO 63121, USA. ¹⁴²Department of Biogeochemical Integration, Max-Planck-Institute for Biogeochemistry, P.O. Box 10 01 64, 07701 Jena, Germany. ¹⁴³Functional Biogeography, Max-Planck-Institute for Biogeochemistry, P.O. Box 10 01 64, 07701 Jena, Germany. ¹⁴⁴Department of Ecology and Evolutionary Biology, UCLA, 621 Charles E. Young Drive South, Box 951606, Los Angeles, CA 90095, USA. ¹⁴⁵Jean-Louis Guillaumet is deceased. ✉email: e.t.pos@uu.nl

Drivers of species distributions and their predictions have been a long-standing search in ecology, with approaches varying from deterministic to neutral (i.e. stochastic) and almost everything in between (e.g. near-neutral, continuum or emergent-neutral^{1,2}). Most models are based on prior assumptions of processes that drive community dynamics. The Maximum Entropy Formalism (hereafter called MEF) makes no such, potentially unjustified, a-priori assumptions in generating predictions of species abundance distributions, as such it is a useful construct to infer processes driving community dynamics given the constraints imposed by prior knowledge (e.g. functional traits or summed regional abundances)³. Quantifying the relative importance of these distinct constraints can thus provide additional answers to understand the complexity of community dynamics (see Supporting Materials SM: boxes S1–S3). This is especially so because, although many different tests are available that link variation in taxon abundances to (1) trait variation, (2) taxon turnover between habitats or environments and (3) the distance decay of similarities between samples, none quantify the importance of these relative to each other. The MEF as applied here, however, is capable of and designed to do exactly this by decomposing variation to separate information explained by each of these aspects in a four-step model (Fig. 1 and Box S2). Its application to an unprecedented large tree inventory database on genus level taxonomy consisting of > 2,000 1-ha plots distributed over Amazonia⁴ and a genus trait database of 13 key functional traits representing global axes of plant strategies⁵ allows us to advance the study of Amazonian tree community dynamics from a new cross-disciplinary perspective.

Results

Principles from information theory^{6–8} can be used in an ecological setting to predict the most likely abundance state for each taxon while simultaneously maximizing entropy based on constraints. Maximization of entropy allows quantifying the information yield for each constraint and therefore identifies which constraints reduce entropy the most. Here we specifically use Shipley's mathematical framework (CATS) for the MEF calculations, similar to earlier studies^{9–11}.

Predictive power of the four-step model. Using a uniform prior and CWM values (Community Weighted Means) as constraints accounted for 23% on average of total deviance between observed and predicted relative abundances (measured by R^2_{KL} values, see Box S2 Eq. 5). Filtered by forest type this was 34% for podzol forests, *várzea* 25%, *igapó* 23%, swamp forests 34%, 21% and 24% for Guyana Shield and Pebas *terra firme* respectively and 20% for Brazilian Shield *terra firme* forests (see Table S1 for detailed decomposition). Using observed metacommunity relative abundances as prior regardless of functional traits accounted for on average 56% for the combined dataset with for all forest types between 51 and 55%, except for the Guyana Shield *terra firme* with 62%. The hybrid model (including both traits as constraints and the metacommunity prior) performed slightly better for the combined dataset (average 60%) with a minimum of 57% for swamp and *várzea* forests and a maximum of 66% for Guyana Shield *terra firme* forests. To compensate for spurious relationships between regional abundances and local trait constraints, regardless of selection, explanatory power was regarded relative to model bias yielding the pure trait and metacommunity effects (Box S3, Fig. 2 and Table S1). This lowered the proportion of information accounted for and yielded average pure metacommunity effects of 40% for the overall dataset ranging between 26 and 45% for each forest type separately with pure trait effects explaining only 5% of information for the combined dataset on average with for each forest type between 3 and 9%. Although the latter was lowered substantially, the explanatory power did appear to be strongly dependent on forest type. The online supplementary material provides additional results relating to the predictive power of each model as well as the spatial gradient of the pure trait and metacommunity effects (Figs. S2–S3).

Direction and strength of selection of trait-based constraints. Each trait showed significant differences in lambda when compared between forest types (Fig. S1, see methods for a definition of lambda). Scatterplots of CWM trait values versus lambda show that, in general, higher lambda values correspond with higher CWM trait values (Fig. S7). A number of functional traits associated with low nutrient conditions (e.g. ectomyorrhiza) and life history strategies suited for protection against herbivores (e.g. latex, resin and high leaf C content) were clearly positively associated with abundance in nutrient poor environments (podzols), indicated by the positive lambda values. In contrast, having fleshy fruits and high leaf N and P content were clearly negatively associated with abundance on these soils. Nodulation was also negatively associated with abundance on poor soils. The ability to accumulate aluminium was positively associated with abundance on those soils commonly associated with higher aluminium content such as *igapó* (strong positive effects). In contrast, it was strongly negatively associated with abundance on other soils, with negative lambda values for podzol and *várzea* forests. Traits such as SLA or winged fruits also showed patterns dependent on forest type, although less pronounced.

Effect of regional metacommunity prior. There was a remarkable similar mean 22% decrease of the information explained purely by the metacommunity prior for each forest type (Fig. S3). For the separate forest types, although the initial pure metacommunity effect varied, the decline appeared remarkably similar with a mean 25% decrease in pure metacommunity effect for podzol, 23% for *várzea*, 26% for *igapó* and 27% for swamp forests with *terra firme* forests having a smaller decline of approximately 21%, averaged over the three subregions. It should be noted there is an obvious risk that when sampling size is increased, this also includes more environmental heterogeneity as samples are coming from a variety of localities potentially leading to changing composition. If this were the case, however, the regional prior (q_i from Fig. 1 and Box S2) would also change, as taxa might be abundant in some places but rare or absent in others. As the metacommunity effect is the explained information that remains relative to any trait effects (i.e. information unique to the neutral prior) and the pure trait effects are the explained information remaining after correcting for pure metacommunity effects

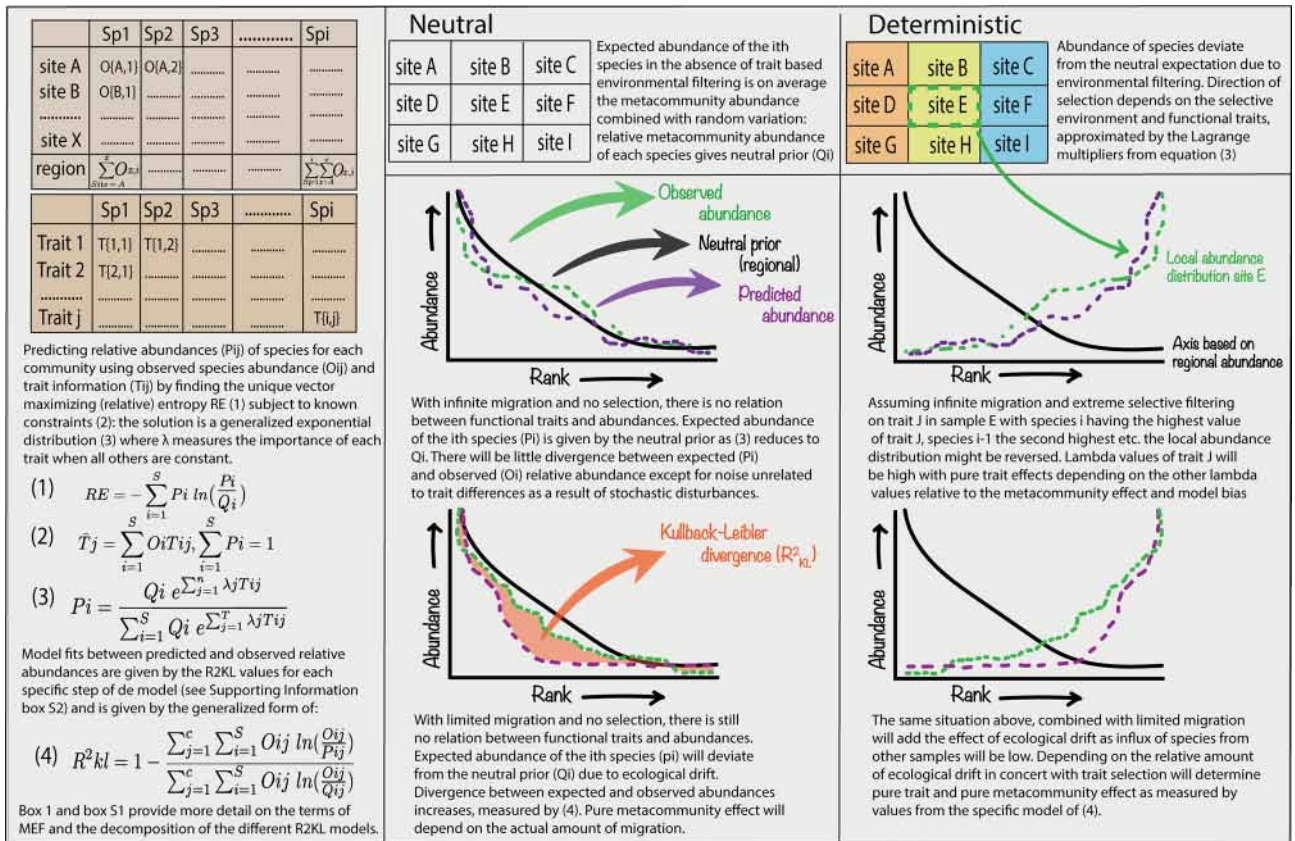


Figure 1. Schematic depiction of the MEF procedure. Left panel shows a genus abundances per site and a functional trait matrix per genus, bottom half outlines calculations. Middle and right panel show different scenarios of neutral and deterministic dynamics under infinite or limited migration. Figure was custom made using Adobe Illustrator (Adobe Inc., 2019. Adobe Illustrator).

(Box S3) this effect should then be accompanied by an increase in pure trait effect for each sample. This was not observed, not even within the different forest types. Instead, the trait effect gradually went up and then remained constant (Fig. S4).

Discussion

The MEF emerges from a well-founded theoretical and empirical body of ecology and evolutionary biology, regarding natural selection, migration and population dynamics. From an ecological point of view, it can be used to quantify the relative association between directional or stabilizing selection for functional traits versus the importance of relative regional abundance regardless of these traits by imposing these as constraints. Our results show that pure trait effects, on average, explained only 5% of the information when all forest types were taken together whereas the pure metacommunity effect, however, explained eight times more taken all forest types together (40%). Greater trait dissimilarity was positively associated with higher pure trait effects, indicating trait-based selection, although the assumed influence of dispersal regardless of these traits appeared to confer more information explaining tree genus composition of the Amazon rainforest. The strength and direction of selection indicated clear directional selective pressure for life history strategies of either growth or protection, depending on forest type (see supplementary online material S-A for a more detailed exploration of ecological interpretation). Including community weighted variance as reflective of potential stabilizing selection did not provide additional information. Although this could be interpreted as indicative of weak or absent stabilizing selection, it is more likely to be an artefact of many genera not being shared among localities due to the sheer geographical scale resulting in a strong mismatch between observed, predicted and uniform relative abundances resulting in a model bias higher than information yielded by including these constraints (see also box S2).

Despite showing clear patterns in environmental selection and dispersal effects, there was a large proportion of information left unexplained (44% on average). Potentially, local demographic stochasticity could weaken any link between functional traits measured and regional abundances of genera. This would, however, mean that almost half of the information contained in relative abundances are the result of random population dynamics and are not structurally governed. Alternatively, this could be due to functional traits reflective of processes not taken into account in this study, such as traits reflective of interactions between trophic levels (e.g. traits influencing specific plant-pathogen interactions). Another and at least equally likely hypothesis for (local) unexplained information is that when scaling up, the ratio of genus richness to total abundance decreases rapidly initially but stabilizes again as relatively non-overlapping habitats are included in the regional abundance distributions

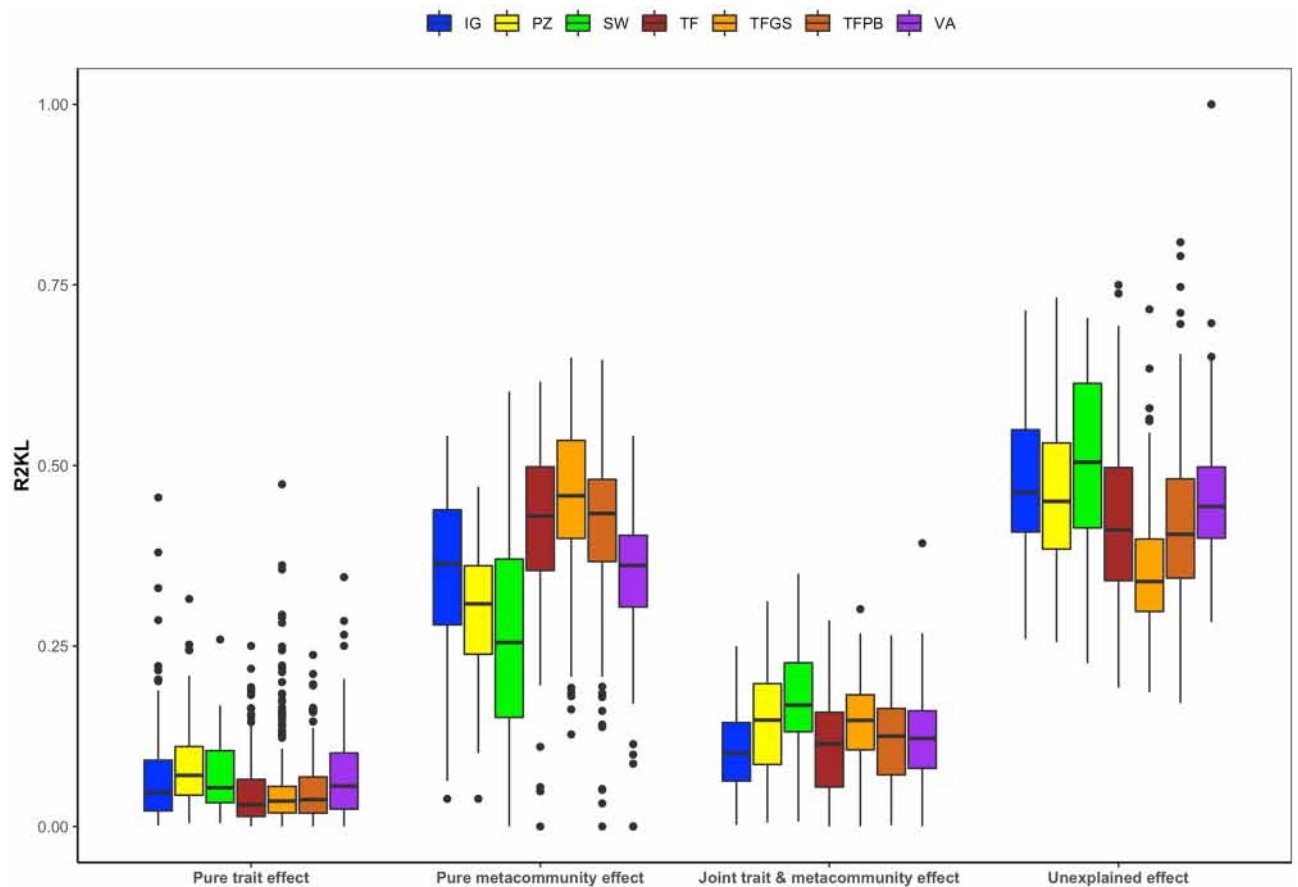


Figure 2. Visual representation of pure trait, pure metacommunity, hybrid model and the remaining unexplained information for each separate forest type. Abbreviations indicate different types: igapó (IG), podzol (PZ), swamp (SW), Brazilian shield terra firme (TFBS), Guiana Shield terra firme (TFGS), Pebas terra firme (TFPB) and várzea (VA). Boxplots show median value of pure effects over all samples, with lower and upper hinges corresponding to 25th and 75th percentiles. Whiskers extends from hinge to largest or smallest value no further than $1.5 \times$ IQR from hinge. Points beyond this range are plotted individually and only positive values were plotted.

and more genera are included again due to the different habitats. This would result in a change of the regional abundance distribution (i.e. the prior) to which each local community is compared, resulting in higher local unexplained information. Further study into these aspects could provide additional insight, though the data necessary for these scales is lacking for Amazonian trees.

Metacommunity importance. Although the initial explanatory power of the metacommunity prior differed between forest types, the decay pattern was very similar. As the effects of either traits or the metacommunity are measured in the goodness-of-fit predictions on local relative abundances, this implies that at small spatial scales the surrounding regional abundances provide better estimators than functional traits, while at larger spatial scales this shifts to the traits. The ecological translation would be that on small spatial scales, local communities share similar environmental conditions leaving dispersal and drift acting predominantly in changing community composition, at least for genus level taxonomy. As the potential regional pool is increased, more and more environmental heterogeneity and non-overlapping regions are likely to be introduced. The more gradual decline of *terra firme* forests can then arguably be attributed to these forests having the largest relative surface area of Amazonia (even for the separate subregions), potentially giving these forests an almost continuous metacommunity without gaps, resulting in a more gradual transition from metacommunity to trait relative importance. The fact that metacommunity effects do not change anymore after certain distances would indicate the effect of dispersal potentially occurs over very large distances. It should be noted that as these calculations are done at community and genus level, they do not measure single dispersal events but rather the effect of dispersal on community composition much deeper in time. In other words, this effect suggests more than a dispersal event every now and then. Instead, it argues for prolonged mixing of forests on large geographical and temporal scales, supported by recent findings demonstrating a lack of geographical phylogenetic structure of lineages for Amazonian tree genera¹².

Conclusion. Using an unprecedented scale of data and applying the Maximum Entropy Formalism from information theory we show that constraints formed by regional relative abundances of genera explain eight times more of local relative abundances than constraints based on directional selection for specific functional traits, although the latter does show clear signals of environmental dependency. There is, however, still much to be explored due to the large unexplained effects and analyses on finer taxonomic (i.e. species level) and environmental (e.g. microhabitat) scales could resolve these issues. The relatively large effects of the regional pool of genera over great distances does suggest an important role for long term dispersal and mixing of Amazonian trees, especially for the Amazonian interior.

Methods

Empirical data. The ATDN^{4,13,14} consists of over 2000 tree inventory plots distributed over the Amazon basin and the Guiana Shield, collectively referred to as Amazonia (a map of all current plots can be found at <https://atdn.myspecies.info/>). Only those plots with trees ≥ 10 cm diameter at breast height were used, leaving 2011 plots with a mean of 558 individuals per plot identified to at least genus level. Most plots used are 1 ha in size (1414) with 492 being smaller (minimum size of 0.1 ha) and 105 larger (maximum size of 80 ha). Genera have been standardized to the W3 Tropicos database¹⁵ using the Taxonomic Name Resolution Service (TNRS, see¹⁶). After filtering based on above criteria and solving nomenclature issues, 1,121,935 individuals belonging to over 828 genera remained. Plots were distributed over seven abiotically different forest types: Podzol forests (PZ), *Igapó* (IG, black water flood forests), *Várzea* (VA white water flood forests), Swamp (SW) and *Terra firme* forests (TF) with subregions BS (Brazilian Shield), GS (Guyana Shield) and PB (Pebas) (see also¹⁷ for details regarding these forest types). Trait data were extracted from several sources. Wood density was mostly derived from¹⁸. Traits related to leaf characteristics mostly came from four large datasets^{19–24}, including additional data from other sources^{25–27} as well as unpublished data (J. Lloyd, A.A. de Oliveira, L. Poorter, M. van de Sande & Mazzei, M. van de Sande & L. Poorter). Data on seed mass came from^{28–30} as well as different floras and tree guides. As this particular trait can vary over several orders of magnitude, this was included on a log-scale^{29,31}. Ectomycorrhizal aspects were derived from literature³², the same applies to nodulation^{33,34}. Traits involved in aluminum accumulation were based on^{35,36} and references therein. For binary traits (yes/no), a genus was considered having a certain trait only when $> 50\%$ of the genus was positive for that specific trait.

Functional traits and trait imputation. Constraints were formed by Community Weighted genus Means (CWM) of functional traits (Table 1), related to key ecological life history aspects. According to principles of natural selection, CWM values will likely be biased towards favourable trait values for that particular environment in the case of directional selection, as taxa with these traits will be more abundant due to environmental selection. Previous studies included community weighted variance (CWV) as well as indicative of potential stabilizing selection^{11,37}. In our case, however, including CWV as constraints resulted in a model bias that was consistently higher than information including trait or metacommunity aspects, CWV was therefore not included as constraints in the final analysis. As for many traits it has been shown earlier that the interspecific variability was larger than the intraspecific variability, this allowed the use of data from different sources to at least calculate a mean species trait value. Genus trait values were subsequently computed as genus-level means of species values if known within the genus and considered constant for each genus. Genus level of taxonomy was used as the available trait database had the most information on this taxonomic level (see Table 1). Unknown values for traits were estimated by Multiple Imputation with Chained Equations (MICE, see¹⁵) by delta adjustment, subtracting a fixed amount (delta), with sensitivity of this adjustment to the imputations of the observed versus imputed data analysed using density plots (Fig. S8) and a linear regression model. This procedure was done using the *mice* package³⁸, available on the R repository, under predictive mean matching (*pmm* setting, 50 iterations). Results showed imputations were stable and showed near identical patterns with each imputation scenario (see Figs. S5–S6 and Table S2). After imputation, all trait values were transformed to Community Weighted Means (CWM) of each trait (*J*) for each plot (*K*) (\bar{T}_{JK}) as $\bar{T}_{JK} = \sum_{i=1}^S t_{ij} ra_{ik}$ with *ra* the relative abundance of the *i*th genus in the *k*th plot following earlier uses³⁷.

MEF procedure predictions and ecological inference. Figure 1 provides a schematic procedure overview, box S1 provides an overview of important terms and Boxes S2–S3 further mathematical details. Initially, a maximally uninformative prior is specified, where q_i (Box S1 Eq. 1) equals $1/S$, indicative of each species having equal abundances, and trait constraints are randomly permuted multiple times ($n=50$) among genera to test whether inclusion of specified constraints significantly changes derived probability distributions (see also³⁹). Subsequently, the same prior is used but now observed trait CWM values belonging to specific genera are used as constraints (following earlier applications using simulated communities¹¹). Third, observed regional abundances are used as prior with permuted trait constraints and finally both observed regional abundances and observed trait CWM are used as prior and constraints. *Maxent2*, an updated version of the *maxent* function currently in the FD library of R provided the computational platform. Proportions of uncertainty explained by each model are given by the Kullback–Leibler divergence R^2_{KL} , a generalization of the classic R^2 goodness of fit. In contrast with standard linear regression models having squared goodness-of-fits measurements, the R^2_{KL} is much more related to the concept of relative entropy, quantifying the information lost when one distribution is compared to another by means of quantifying the statistical distance between two distributions⁴⁰. Pure trait, pure metacommunity, joint metacommunity-trait and unexplained effects are calculated as proportions of total biologically relevant information (Box S1 and Box S2). Data was rarefied to smallest sample size (swamp forests; 28) and calculations bootstrapped 25 times. Results indicated no significant change compared to using all data, hence the total dataset was used for all analyses.

| Functional trait | Units | Mean | SD | Est % | Associated challenge |
|--|---------------------|------|------|-------|---|
| Wood density (<i>WD</i>) | g/cm ³ | 0.63 | 0.17 | 30 | Longevity ⁴³ |
| Seed mass class (<i>SMC</i>) | categorical (1–8) | 4.3 | 1.4 | 31 | Dispersal, Fecundity, Establishment ⁴³ |
| Specific leaf area (<i>SLA</i>) | mm ² /mg | 15 | 5.9 | 41 | Establishment, Plasticity, Disturbance ⁴³ |
| Leaf nitrogen content (<i>N</i>) | mg/g | 22.3 | 7.30 | 41 | Photosynthetic capacity ⁴³ |
| Leaf phosphorus content (<i>P</i>) | mg/g | 1 | 0.77 | 50 | Limited available P for metabolism ⁴⁴ |
| Leaf carbon content (<i>C</i>) | mg/g | 468 | 38.1 | 54 | Herbivore resistance (C:N) ⁴⁵ |
| Latex | 1 = no, 2 = yes | 1.2 | 0.43 | 46 | Herbivore resistance ⁴⁶ |
| Resin | 1 = no, 2 = yes | 1.1 | 0.35 | 58 | Herbivore resistance ⁴⁶ |
| Root nodules (<i>Nodules</i>) | 1 = no, 2 = yes | 1.1 | 0.28 | 0 | Nitrogen fixation ⁴⁷ |
| Ectomycorrhiza (<i>EctoMyco</i>) | 1 = no, 2 = yes | 1.01 | 0.11 | 0 | Organic N fixation, heavy metal pollution ⁴⁸ |
| Aluminum accumulation (<i>AlAcc</i>) | 1 = no, 2 = yes | 1.1 | 0.21 | 3 | Heavy metal pollution ⁴⁹ |
| Fleshy fruits (<i>Fleshy</i>) | 1 = no, 2 = yes | 1.6 | 0.50 | 7 | Dispersal (<i>specificity</i>) ⁵⁰ |
| Winged seeds (<i>Wings</i>) | 1 = no, 2 = yes | 1.2 | 0.42 | 39 | Dispersal (<i>limitation</i>) ⁵⁰ |

Table 1. Overview of used functional traits. Mean and standard deviation (SD) are calculated after predictive mean matching (percentage of estimated values is given by Est (%)). Associated challenge indicates different aspects of life history and selective environment related to specific functional traits, sources are given in the footnote. For specific methodology of measurement protocols and calculation for each trait we refer to the original sources of the data (see main text).

Strength and direction of selection. Predictions of genus relative abundances are computed as a function of traits reflected in the CWM values and a series of constants (λ_{jk} : the Lagrange Multipliers). Each multiplier quantifies the association between a unit of change for a particular trait j and a proportional change in predicted relative abundance p_{ik} (the i th genus in the k th community) considering all other traits are constant, formally described as: $\partial p_{ik} / \partial t_{ij} = \lambda_{jk} p_{ik} (1 - p_{ik})$ (see appendix 1 from⁴¹). Positive values indicate larger trait values associated with higher abundances (positive selection), negative values indicate the opposite (negative selection) with changes proportional to lambda. Values approximating zero indicate no association between specific traits and relative abundances of species. Decomposing λ_{jk} and comparing by means of a One-Way Analysis of Variance for each trait separately between forest types allows studying both the strength and direction of selection in different habitats. Note that this is done for the same constraint between forest types, as lambda values for each constraint do not scale linearly between different constraints.

Estimation of metacommunity size. Iteratively increasing the regional species pool considered as prior in concentric circles of a fixed radius of 50 km allows estimating the spatial effect of metacommunity size. Due to computational limits, the number of permutations for the MEF calculations (see above) was reduced to two, shuffling the combinations of genera and traits at least once. Comparison of results from the analyses using all plots indicated small effects of a smaller perturbations (average of 5% difference for metacommunity effect between 5 and 50 permutations). The relationship between pure metacommunity effect and radius of metacommunity size was fitted using a smoothing loess regression (function *loess* and *predict*; R-package *stats* with span set at 0.1). Fits subsequently were used to predict values of metacommunity effect based on geographical distance to visualize general patterns for each forest type. Exponential decay of pure metacommunity effect was described using a self-start asymptotic regression function (*SSasymp*) of the form $y(t) \sim y_f + (y_0 - y_f)e^{-\exp(\log(\alpha))t}$ (*nls* from *stats*⁴²). A list of all packages used in R in addition to those preloaded can be found in the supplementary online material (SA2).

Data availability

R scripts are available on the github repository of E.T. Pos (EdwinTPos). The data that support the findings of this study are available from The Amazon Tree Diversity Network (ATDN) upon reasonable request.

Received: 11 March 2021; Accepted: 13 January 2023

Published online: 17 February 2023

References

- Chisholm, R. A. & Pacala, S. W. Niche and neutral models predict asymptotically equivalent species abundance distributions in high-diversity ecological communities. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 15821 (2010).
- Tilman, D. Niche tradeoffs, neutrality, and community structure: A stochastic theory of resource competition, invasion, and community assembly. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 10854 (2004).
- Harte, J. *Maximum Entropy and Ecology: A Theory of Abundance, Distribution, and Energetics* (OUP Oxford, 2011).
- The Amazon Tree Diversity Network, <http://atdn.myspecies.info/>.
- Díaz, S. *et al.* The global spectrum of plant form and function. *Nature* **529**, 167 (2016).
- Jaynes, E. T. Information theory and statistical mechanics I. *Phys. Rev.* **106**, 620 (1957).
- Jaynes, E. T. Information theory and statistical mechanics. II. *Phys. Rev.* **108**, 171 (1957).
- Shannon, C. E. A mathematical theory of communication. *Bell Syst. Tech. J.* **27**, 379 (1948).

9. Shipley, B. Measuring and interpreting trait-based selection versus meta-community effects during local community assembly. *J. Veg. Sci.* **25**, 55 (2014).
10. Shipley, B., Paine, C. E. T. & Baraloto, C. Quantifying the importance of local niche-based and stochastic processes to tropical tree community assembly. *Ecology* **93**, 760 (2012).
11. Loranger, J., Munoz, F., Shipley, B. & Violle, C. What makes trait-abundance relationships when both environmental filtering and stochastic neutral dynamics are at play?. *Oikos* **127**, 1735 (2018).
12. Dexter, K. G. *et al.* Dispersal assembly of rain forest tree communities across the amazon basin. *Proc. Natl. Acad. Sci.* **114**, 2645 (2017).
13. ter Steege, H. *et al.* Hyperdominance in the amazonian tree flora. *Science* **342**, 1243092 (2013).
14. ter Steege, H. *et al.* Biased-corrected richness estimates for the amazonian tree flora. *Sci. Rep.* **10**, 1 (2020).
15. *Tropicos, Tropicos.Org. Missouri Botanical Garden. Accessed 2018, Version Number Unknown Due to Lack of Release-History.*
16. Boyle, B. *et al.* The taxonomic name resolution service: An online tool for automated standardization of plant names. *BMC Bioinform.* **14**, 16 (2013).
17. Haugaasen, T. & Peres, C. Floristic, edaphic and structural characteristics of flooded and unflooded forests in the lower rio purús region of Central Amazonia, Brazil. *Acta Amaz.* **36**, 25 (2006).
18. Chave, J., Coomes, D. & Jansen, S. Towards a worldwide wood economics spectrum. *Ecol. Lett.* **351**, 855 (2009).
19. Kraft, N. J. B., Valencia, R. & Ackerly, D. D. Functional traits and niche-based tree community assembly in an amazonian forest. *Science* **322**, 580 (2008).
20. Fyllas, N. M. *et al.* Basin-wide variations in foliar properties of amazonian forest: Phylogeny, soils and climate. *Biogeosciences* **6**, 2677 (2009).
21. Baraloto, C. *et al.* Decoupled leaf and stem economics in rain forest trees. *Ecol. Lett.* **13**, 1338 (2010).
22. Fortunel, C., Fine, P. V. A. & Baraloto, C. Leaf, stem and root tissue strategies across 758 neotropical tree species. *Funct. Ecol.* **26**, 1153 (2012).
23. Paine, C. E. T., Baraloto, C., Chave, J. & Hérault, B. Functional traits of individual trees reveal ecological constraints on community assembly in tropical rain forests. *Oikos* **120**, 720 (2011).
24. Patiño, S. *et al.* Coordination of physiological and structural traits in amazon forest trees. *Biogeosciences* **9**, 775 (2012).
25. F. M. J. Ohler, *Phytomass and Mineral Content in Untouched Forest*, (1980).
26. Thompson, J. *et al.* Ecological studies on a lowland evergreen rain forest on Maraca Island, Roraima, Brazil. I. Physical environment, forest structure and leaf chemistry. *J. Ecol.* **689**, 89 (1992).
27. Pons, T. L., Perreijn, K., Van Kessel, C. & Werger, M. J. A. Symbiotic nitrogen fixation in a tropical rainforest: 15N natural abundance measurements supported by experimental isotopic enrichment. *New Phytol.* **173**, 154 (2006).
28. Foster, S. & Janson, C. H. The relationship between seed size and establishment conditions in tropical woody plants. *Ecology* **66**, 773 (1985).
29. Hammond, D. S. & Brown, V. K. Hammond and brown 1995 seed size. Ecology.Pdf. *Ecology* **76**, 2544 (1995).
30. Royal Botanic Gardens Kew, 2020, *Seed Information Database (SID)*.
31. ter Steege, H. & Hammond, D. Character convergence, diversity, and disturbance in tropical rain forest in Guyana. *Ecology* **82**, 3197 (2001).
32. Tedersoo, L. & Brundrett, M. C. Evolution of ectomycorrhizal symbiosis in plants. *Biogeogr. Mycorrhizal Symbiosis*. **407**, 89 (2017).
33. Sprent, J. I. *Nodulation in Legumes*, Royal Botanic Gardens, Kew.
34. Soltis, P. S., Soltis, D. E. & Chase, M. W. Angiosperm phylogeny inferred from multiple genes as a tool for comparative biology. *Nature* **402**, 402 (1999).
35. Jansen, S., Watanabe, T., Dessein, S., Smets, E. & Robbrecht, E. A comparative study of metal levels in leaves of some Al-accumulating *Rubiaceae*. *Ann. Bot.* **91**, 657 (2003).
36. Jansen, S., Watanabe, T. & Smets, E. Aluminium accumulation in leaves of 127 species in *Melastomataceae*, with comments on the order myrtales. *Ann. Bot.* **90**, 53 (2002).
37. Sonnier, G., Shipley, B. & Navas, M. L. Quantifying relationships between traits and explicitly measured gradients of stress and disturbance in early successional plant communities. *J. Veg. Sci.* **21**, 1014 (2010).
38. van Buuren, S. & Groothuis-Oudshoorn, K. Mice: Multivariate imputation by chained equations in R. *J. Stat. Softw.* **45**, 1 (2011).
39. Roxburgh, S. H. & Mokany, K. On testing predictions of species relative abundance from maximum entropy optimisation. *Oikos* **119**, 583 (2010).
40. Cameron, A. C. & Windmeijer, F. A. G. An R-squared measure of goodness of fit for some common nonlinear regression models. *J. Econ.* **77**, 329 (1997).
41. Sonnier, G., Navas, M. L., Fayolle, A. & Shipley, B. Quantifying trait selection driving community assembly: A test in herbaceous plant communities under contrasted land use regimes. *Oikos* **121**, 1103 (2012).
42. R Core Team, *R: A Language and Environment for Statistical Computing*.
43. Weiher, E. *et al.* Challenging theophrastus: A common core list of plant traits for functional ecology. *J. Veg. Sci.* **10**, 609 (1999).
44. Raghothama, K. G. Phosphate acquisition. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **50**, 665 (1999).
45. Poorter, L., Van De Plassche, M., Willems, S. & Boot, R. G. A. Leaf traits and herbivory rates of tropical tree species differing in successional status. *Plant Biol.* **6**, 746 (2004).
46. Agrawal, A. A. Macroevolution of plant defense strategies. *Trends Ecol. Evol.* **22**, 103 (2007).
47. Sprent, J. I. & Parsons, R. Nitrogen fixation in legume and non-legume trees. *F. Crop. Res.* **65**, 183 (2000).
48. Colpaert, J. V., Wevers, J. H. L., Krzmaric, E. & Adriaensen, K. How metal-tolerant ecotypes of ectomycorrhizal fungi protect plants from heavy metal pollution. *Ann. For. Sci.* **68**, 17 (2011).
49. Poschenrieder, C., Günsé, B., Corrales, I. & Barceló, J. A. Glance into aluminum toxicity and resistance in plants. *Sci. Total Environ.* **400**, 356 (2008).
50. Howe, H. F. & Smallwood, J. Ecology of Seed Dispersal. *Ecology* **13**, 201 (2013).

Acknowledgements

This paper is the result of the work of hundreds of different scientists and research institutions in the Amazon over the past 80 years. Without their hard work this analysis would have been impossible. The 25-ha Long-Term Ecological Research Project of Amacayacu is a collaborative project of the Instituto Amazónico de Investigaciones Científicas Sinchi and the Universidad Nacional de Colombia Sede Medellín, in partnership with the Unidad de Manejo Especial de Parques Naturales Nacionales and the Center for Tropical Forest Science of the Smithsonian Tropical Research Institute (CTFS). The Amacayacu Forest Dynamics Plot is part of the Center for Tropical Forest Science, a global network of large-scale demographic tree plots. We acknowledge the Director and staff of the Amacayacu National Park for supporting and maintaining the project in this National Park. We would also like to thank Prof. John Harte, Prof. Bill Shipley and the anonymous reviewers for their constructive feedback during the review process.

Author contributions

E.T.P. and H.T.S. designed the study. E.T.P. performed analyses and took the lead in writing the manuscript, H.T.S. supervised the writing and provided regular feedback both for the manuscript and the interpretation of the results. All other authors provided feedback on the manuscript and provided their data from the Amazon Tree Diversity Network or trait data. Authors E.T.P. to L.V.G. provided tree inventory data, authors G.B., J.K., N.K., A.L., K.M., G.P., L.P. provided data on functional traits, C.B., J.L., A.A.O. and H.T.S. provided both tree inventory and functional trait data.

Funding

HtS and RS were supported by grant 407232/2013-3 – PVE - MEC/MCTI/CAPES/CNPq/FAPs; CB was supported by grant FAPESP 95/3058-0 - CRS 068/96 WWF Brasil - The Body Shop; DS, JFM, JE, PP and JC benefited from an “Investissement d’Avenir” grant managed by the Agence Nationale de la Recherche (CEBA: ANR-10-LABX-25-01); HLQ/MAP/JLLM received financial support by MCT/CNPq/CT-INFRA/GEOMA #550373/2010-1 and # 457515/2012-0, and JLLM were supported by grant CAPES/PDSE # 88881.135761/2016-01 and CAPES/Fapespa #1530801; The Brazilian National Research Council (CNPq) provided a productivity grant to EMV (Grant 308040/2017-1); Floristic identification in plots in the RAINFOR forest monitoring network have been supported by the Natural Environment Research Council (grants NE/B503384/1, NE/ D01025X/1, NE/I02982X/1, NE/F005806/1, NE/D005590/1 and NE/I028122/1) and the Gordon and Betty Moore Foundation; B.M.F. is funded by FAPESP grant 2016/25086-3. BSM, BHMJ and OLP were supported by grants CNPq/CAPES/FAPS/BC-Newton Fund #441244/2016-5 and FAPEMAT/0589267/2016; TWH was funded by National Science Foundation grant DEB-1556338.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-023-28132-y>.

Correspondence and requests for materials should be addressed to E.P.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023