

1 **Title:** The Maximum Entropy Formalism of statistical mechanics in a biological
2 application: a quantitative analysis of tropical forest ecology.

3
4 **Authors:** Edwin Pos*,1,2 Luiz de Souza Coelho,3 Diogenes de Andrade Lima
5 Filho,3 Rafael P. Salomão,4,5 Iêda Leão Amaral,3 Francisca Dionízia de Almeida
6 Matos,3 Carolina V. Castilho,6 Oliver L. Phillips,7 Juan Ernesto Guevara,8,9 Marcelo
7 de Jesus Veiga Carim,10 Dairon Cárdenas López,11 William E. Magnusson,12
8 Florian Wittmann,13,14 Mariana Victória Irume,3 Maria Pires Martins,3 Daniel
9 Sabatier,15 José Renan da Silva Guimarães,10 Jean-François Molino,15 Olaf S.
10 Bánki,2 Maria Teresa Fernandez Piedade,16 Nigel C.A. Pitman,17 Abel Monteagudo
11 Mendoza,18 José Ferreira Ramos,3 Joseph E. Hawes,19 Everton José Almeida,20
12 Luciane Ferreira Barbosa,20 Larissa Cavalheiro,20 Márcia Cléia Vilela dos Santos,20
13 Bruno Garcia Luize,21 Evlyn Márcia Moraes de Leão Novo,22 Percy Núñez
14 Vargas,23 Thiago Sanna Freire Silva,24 Eduardo Martins Venticinque,25 Angelo
15 Gilberto Manzatto,26 Neidiane Farias Costa Reis,27 John Terborgh,28,29 Katia
16 Regina Casula,27 Euridice N. Honorio Coronado,30,7 Juan Carlos Montero,31,3
17 Beatriz S. Marimon,32 Ben Hur Marimon-Junior,32 Ted R. Feldpausch,33,7 Alvaro
18 Duque,34 Chris Baraloto,35 Nicolás Castaño Arboleda,11 Julien Engel,15,35 Pascal
19 Petronelli,36 Charles Eugene Zartman,3 Timothy J. Killeen,37 Rodolfo Vasquez,18
20 Bonifacio Mostacedo,38 Rafael L. Assis,39 Jochen Schöngart,16 Hernán
21 Castellanos,40 Marcelo Brilhante de Medeiros,41 Marcelo Fragomeni Simon,41 Ana
22 Andrade,42 José Luís Camargo,42 Layon O. Demarchi,16 William F. Laurance,29
23 Susan G.W. Laurance,29 Emanuelle de Sousa Farias,43,44 Maria Aparecida Lopes,45
24 José Leonardo Lima Magalhães,46,47 Henrique Eduardo Mendonça Nascimento,3
25 Helder Lima de Queiroz,48 Gerardo A. Aymard C.,49 Roel Brienem,7 Juan David
26 Cardenas Revilla,3 Flávia R.C. Costa,3 Adriano Quaresma,16 Ima Célia Guimarães
27 Vieira,5 Bruno Barçante Ladvocat Cintra,50 Pablo R. Stevenson,51 Yuri Oliveira
28 Feitosa,52 Joost F. Duivenvoorden,53 Hugo F. Mogollón,54 Leandro Valle Ferreira,5
29 James A. Comiskey,55,56 Freddie Draper,57,35 José Julio de Toledo,58 Gabriel
30 Damasco,59 Nállarett Dávila,60 Roosevelt García-Villacorta,61,62 Aline Lopes,16,63
31 Alberto Vicentini,12 Janaína Costa Noronha,64 Flávia Rodrigues Barbosa,64
32 Rainiellen de Sá Carpanedo,64 Thaise Emilio,65,12 Carolina Levis,66,67 Domingos
33 de Jesus Rodrigues,64 Juliana Schietti,3 Priscila Souza,3 Alfonso Alonso,56
34 Francisco Dallmeier,56 Vitor H.F. Gomes,68,69 Jon Lloyd,70 David Neill,71 Daniel
35 Praia Portela de Aguiar,16 Alejandro Araujo-Murakami,72 Luzmila Arroyo,72
36 Fernanda Antunes Carvalho,12,73 Fernanda Coelho de Souza,12,7 Dário Dantas do
37 Amaral,5 Kenneth J. Feeley,74,75 Rogerio Gribel,3 Marcelo Petratti Pansonato,3,76
38 Jos Barlow,77 Erika Berenguer,78,77 Joice Ferreira,47 Paul V.A. Fine,59 Marcelino
39 Carneiro Guedes,79 Eliana M. Jimenez,80 Juan Carlos Licona,31 Maria Cristina
40 Peñuela Mora,81 Carlos A. Peres,82 Boris Eduardo Villa Zegarra,83 Carlos Cerón,84
41 Terry W. Henkel,85 Paul Maas,2 Marcos Silveira,86 Juliana Stropp,87 Raquel
42 Thomas-Caesar,88 Tim R. Baker,7 Doug Daly,89 Kyle G. Dexter,90,91 John Ethan
43 Householder,13 Isau Huamantupa-Chuquimaco,23 Toby Pennington,33,91 Marcos
44 Ríos Paredes,92 Alfredo Fuentes,93,94 José Luis Marcelo Pena,95 Miles R.
45 Silman,96 J. Sebastián Tello,94 Jerome Chave,97 Fernando Cornejo Valverde,98
46 Anthony Di Fiore,99 Renato Richard Hilário,58 Juan Fernando Phillips,100 Gonzalo
47 Rivas-Torres,101,102 Tinde R. van Andel,2,103 Patricio von Hildebrand,104
48 Edelcilio Marques Barbosa,3 Luiz Carlos de Matos Bonates,3 Hilda Paulette Dávila

49 Doza,92 Émile Fonty,105,15 Ricardo Zárata Gómez,106 Therany Gonzales,107
50 George Pepe Gallardo Gonzales,92 Jean-Louis Guillaumet†,108 Bruce Hoffman,109
51 André Braga Junqueira,110 Yadvinder Malhi,111 Ires Paula de Andrade Miranda,3
52 Linder Felipe Mozombite Pinto,92 Adriana Prieto,112 Agustín Rudas,112 Ademir R.
53 Ruschel,47 Natalino Silva,113 César I.A. Vela,114 Vincent Antoine Vos,115 Eglée L.
54 Zent,116 Stanford Zent,116 Bianca Weiss Albuquerque,16 Angela Cano,51,117
55 Diego F. Correa,51,118 Janaina Barbosa Pedrosa Costa,79 Bernardo Monteiro
56 Flores,119 Milena Holmgren,120 Marcelo Trindade Nascimento,121 Alexandre A.
57 Oliveira,76 Hirma Ramirez-Angulo,122 Maira Rocha,16 Veridiana Vizoni
58 Scudeller,123 Rodrigo Sierra,124 Milton Tirado,124 Maria Natalia Umaña,125
59 Geertje van der Heijden,126 Emilio Vilanova Torre,122,127 Corine Vriesendorp,17
60 Ophelia Wang,128 Kenneth R. Young,129 Manuel Augusto Ahuite Reategui,130
61 Cláudia Baider,131,76 Henrik Balslev,132 Sasha Cárdenas,51 Luisa Fernanda
62 Casas,51 William Farfan-Rios,23,133,94 Cid Ferreira,3 Reynaldo Linares-
63 Palomino,56 Casimiro Mendoza,134,135 Italo Mesones,59 Armando Torres-
64 Lezama,122 Ligia Estela Urrego Giraldo,34 Daniel Villarroel,72 Roderick Zagt,136
65 Miguel N. Alexiades,137 Karina Garcia-Cabrera,96 Lionel Hernandez,40 William
66 Milliken,65 Walter Palacios Cuenca,138 Susamar Pansini,27 Daniela Pauletto,139
67 Freddy Ramirez Arevalo,140 Adeilza Felipe Sampaio,27 Elvis H. Valderrama
68 Sandoval,141,140 Luis Valenzuela Gamarra,18 Gerhard Boenisch,142 Jens
69 Kattge,143 Nathan Kraft,144 Aurora Levesley,7 Karina Melgaço,7 Georgia
70 Pickavance,7 Lourens Poorter,67 Hans ter Steege,2,145

71

72 * Corresponding author: e.t.pos@uu.nl

73 † Deceased 01-2018

74

75 Affiliations

76 1Ecology & Biodiversity Group, Utrecht University, Padualaan 8, Utrecht, 3584 CH,
77 The Netherlands

78 2Naturalis Biodiversity Center, PO Box 9517, Leiden, 2300 RA, The Netherlands

79 3Coordenação de Biodiversidade, Instituto Nacional de Pesquisas da Amazônia -
80 INPA, Av. André Araújo, 2936, Petrópolis, Manaus, AM, 69067-375, Brazil

81 4Programa Professor Visitante Nacional Sênior na Amazônia - CAPES, Universidade
82 Federal Rural da Amazônia, Av. Perimetral, s/n, Belém, PA, Brazil

83 5Coordenação de Botânica, Museu Paraense Emílio Goeldi, Av. Magalhães Barata
84 376, C.P. 399, Belém, PA, 66040-170, Brazil

85 6EMBRAPA – Centro de Pesquisa Agroflorestal de Roraima, BR 174, km 8 –
86 Distrito Industrial, Boa Vista, RR, 69301-970, Brazil

87 7School of Geography, University of Leeds, Woodhouse Lane, Leeds, LS2 9JT, UK

88 8Grupo de Investigación en Biodiversidad, Medio Ambiente y Salud-BIOMAS,
89 Universidad de las Américas, Campus Queri, Quito, Ecuador

90 9Keller Science Action Center, The Field Museum, 1400 S. Lake Shore Drive,
91 Chicago, IL, 60605-2496, USA

92 10Departamento de Botânica, Instituto de Pesquisas Científicas e Tecnológicas do
93 Amapá - IEPA, Rodovia JK, Km 10, Campus do IEPA da Fazendinha, Amapá,
94 68901-025, Brazil

95 11Herbario Amazónico Colombiano, Instituto SINCHI, Calle 20 No 5-44, Bogotá,
96 DC, Colombia

- 97 12Coordenação de Pesquisas em Ecologia, Instituto Nacional de Pesquisas da
98 Amazônia - INPA, Av. André Araújo, 2936, Petrópolis, Manaus, AM, 69067-375,
99 Brazil
- 100 13Dep. of Wetland Ecology, Institute of Geography and Geoecology, Karlsruhe
101 Institute of Technology - KIT, Josefstr.1, Rastatt, D-76437, Germany
- 102 14Biogeochemistry, Max Planck Institute for Chemistry, Hahn-Meitner Weg 1,
103 Mainz, 55128, Germany
- 104 15AMAP, IRD, Cirad, CNRS, INRA, Université de Montpellier, Montpellier, F-
105 34398, France
- 106 16Coordenação de Dinâmica Ambiental, Instituto Nacional de Pesquisas da
107 Amazônia - INPA, Av. André Araújo, 2936, Petrópolis, Manaus, AM, 69067-375,
108 Brazil
- 109 17Science and Education, The Field Museum, 1400 S. Lake Shore Drive, Chicago,
110 IL, 60605-2496, USA
- 111 18Jardín Botánico de Missouri, Oxapampa, Pasco, Peru
- 112 19Applied Ecology Research Group, School of Life Sciences, Anglia Ruskin
113 University, East Road, Cambridge, CB1 1PT, UK
- 114 20ICNHS, Universidade Federal de Mato Grosso, Av. Alexandre Ferronato, 1200,
115 Sinop, MT, 78557-267, Brazil
- 116 21Departamento de Ecologia, Universidade Estadual Paulista - UNESP – Instituto de
117 Biociências – IB, Av. 24 A, 1515, Bela Vista, Rio Claro, SP, 13506-900, Brazil
- 118 22Divisao de Sensoriamento Remoto – DSR, Instituto Nacional de Pesquisas
119 Espaciais – INPE, Av. dos Astronautas, 1758, Jardim da Granja, São José dos
120 Campos, SP, 12227-010, Brazil
- 121 23Herbario Vargas, Universidad Nacional de San Antonio Abad del Cusco, Avenida
122 de la Cultura, Nro 733, Cusco, Cuzco, Peru
- 123 24Biological and Environmental Sciences, University of Stirling, Stirling, FK9 4LA,
124 UK
- 125 25Centro de Biociências, Departamento de Ecologia, Universidade Federal do Rio
126 Grande do Norte, Av. Senador Salgado Filho, 3000 , Natal, RN, 59072-970, Brazil
- 127 26Departamento de Biologia, Universidade Federal de Rondônia, Rodovia BR 364 s/n
128 Km 9,5 - Sentido Acre, Unir, Porto Velho, RO, 76.824-027, Brazil
- 129 27Programa de Pós- Graduação em Biodiversidade e Biotecnologia PPG- Bionorte,
130 Universidade Federal de Rondônia, Campus Porto Velho Km 9,5 bairro Rural, Porto
131 Velho, RO, 76.824-027, Brazil
- 132 28Department of Biology and Florida Museum of Natural History, University of
133 Florida, Gainesville, FL, 32611, USA
- 134 29Centre for Tropical Environmental and Sustainability Science and College of
135 Science and Engineering, James Cook University, Cairns, Queensland, 4870,
136 Australia
- 137 30Instituto de Investigaciones de la Amazonía Peruana (IIAP), Av. A. Quiñones km
138 2,5, Iquitos, Loreto, 784, Peru
- 139 31Instituto Boliviano de Investigacion Forestal, Av. 6 de agosto #28, Km. 14, Doble
140 via La Guardia, Casilla 6204, Santa Cruz, Santa Cruz, Bolivia
- 141 32Programa de Pós-Graduação em Ecologia e Conservação, Universidade do Estado
142 de Mato Grosso, Nova Xavantina, MT, Brazil
- 143 33Geography, College of Life and Environmental Sciences, University of Exeter,
144 Rennes Drive, Exeter, EX4 4RJ, UK
- 145 34Departamento de Ciencias Forestales, Universidad Nacional de Colombia, Calle 64
146 x Cra 65, Medellín, Antioquia, 1027, Colombia

- 147 35International Center for Tropical Botany (ICTB) Department of Biological
148 Sciences, Florida International University, 11200 SW 8th Street, OE 243, Miami, FL,
149 33199, USA
150 36Cirad UMR Ecofog, AgrosParisTech,CNRS,INRA,Univ Guyane, Campus
151 agronomique, Kourou Cedex, 97379, France
152 37Agteca-Amazonica, Santa Cruz, Bolivia
153 38Facultad de Ciencias Agrícolas, Universidad Autónoma Gabriel René Moreno,
154 Santa Cruz, Santa Cruz, Bolivia
155 39Natural History Museum, University of Oslo, Postboks 1172, Oslo, 0318, Norway
156 40Centro de Investigaciones Ecológicas de Guayana, Universidad Nacional
157 Experimental de Guayana, Calle Chile, urbaniz Chilemex, Puerto Ordaz, Bolivar,
158 Venezuela
159 41Prédio da Botânica e Ecologia, Embrapa Recursos Genéticos e Biotecnologia,
160 Parque Estação Biológica, Av. W5 Norte, Brasilia, DF, 70770-917, Brazil
161 42Projeto Dinâmica Biológica de Fragmentos Florestais, Instituto Nacional de
162 Pesquisas da Amazônia - INPA, Av. André Araújo, 2936, Petrópolis, Manaus, AM,
163 69067-375, Brazil
164 43Laboratório de Ecologia de Doenças Transmissíveis da Amazônia (EDTA),
165 Instituto Leônidas e Maria Deane, Fiocruz, Rua Terezina, 476, Adrianópolis, Manaus,
166 AM, 69060-001, Brazil
167 44Programa de Pós-graduação em Biodiversidade e Saúde, Instituto Oswaldo Cruz -
168 IOC/FIOCRUZ, Pav. Arthur Neiva – Térreo, Av. Brasil, 4365 – Manguinhos, Rio de
169 Janeiro, RJ, 21040-360, Brazil
170 45Instituto de Ciências Biológicas, Universidade Federal do Pará, Av. Augusto
171 Corrêa 01, Belém, PA, 66075-110, Brazil
172 46Programa de Pós-Graduação em Ecologia, Universidade Federal do Pará, Av.
173 Augusto Corrêa 01, Belém, PA, 66075-110, Brazil
174 47Embrapa Amazônia Oriental, Trav. Dr. Enéas Pinheiro s/nº, Belém, PA, 66095-
175 100, Brazil
176 48Diretoria Técnico-Científica, Instituto de Desenvolvimento Sustentável Mamirauá,
177 Estrada do Bexiga, 2584, Tefé, AM, 69470-000, Brazil
178 49Programa de Ciencias del Agro y el Mar, Herbario Universitario (PORT),
179 UNELLEZ-Guanare, Guanare, Portuguesa, 3350, Venezuela
180 50Instituto de Biociências - Dept. Botanica, Universidade de Sao Paulo - USP, Rua
181 do Matão 277, Cidade Universitária, São Paulo, SP, 05508-090, Brazil
182 51Laboratorio de Ecología de Bosques Tropicales y Primatología, Universidad de los
183 Andes, Carrera 1 # 18a- 10, Bogotá, DC, 111711, Colombia
184 52Programa de Pós-Graduação em Biologia (Botânica), Instituto Nacional de
185 Pesquisas da Amazônia - INPA, Av. André Araújo, 2936, Petrópolis, Manaus, AM,
186 69067-375, Brazil
187 53Institute of Biodiversity and Ecosystem Dynamics, University of Amsterdam,
188 Sciencepark 904, Amsterdam, 1098 XH, The Netherlands
189 54Endangered Species Coalition, 8530 Geren Rd., Silver Spring, MD, 20901, USA
190 55Inventory and Monitoring Program, National Park Service, 120 Chatham Lane,
191 Fredericksburg, VA, 22405, USA
192 56Center for Conservation and Sustainability, Smithsonian Conservation Biology
193 Institute, 1100 Jefferson Dr. SW, Suite 3123, Washington, DC, 20560-0705, USA
194 57Department of Global Ecology, Carnegie Institution for Science, 260 Panama St.,
195 Stanford, CA, 94305, USA

- 196 58Universidade Federal do Amapá, Ciências Ambientais, Rod. Juscelino Kubitschek
197 km2, Macapá, AP, 68902-280, Brazil
- 198 59Department of Integrative Biology, University of California, Berkeley, CA, 94720-
199 3140, USA
- 200 60Biologia Vegetal, Universidade Estadual de Campinas, Caixa Postal 6109,
201 Campinas, SP, 13.083-970, Brazil
- 202 61Department of Ecology and Evolutionary Biology, Cornell University, Corson Hall,
203 215 Tower Road, Ithaca, NY, 14850, USA
- 204 62Peruvian Center for Biodiversity and Conservation (PCBC), Iquitos, Peru
- 205 63Department of Ecology, University of Brasilia, Brasilia, DF, 70904-970, Brazil
- 206 64ICNHS, Federal University of Mato Grosso, Av. Alexandre Ferronato 1200, Setor
207 Industrial, Sinop, MT, 78.557-267, Brazil
- 208 65Natural Capital and Plant Health, Royal Botanic Gardens, Kew, Richmond, Surrey,
209 TW9 3AB, UK
- 210 66Programa de Pós-Graduação em Ecologia, Instituto Nacional de Pesquisas da
211 Amazônia - INPA, Av. André Araújo, 2936, Petrópolis, Manaus, AM, 69067-375,
212 Brazil
- 213 67Forest Ecology and Forest Management Group, Wageningen University &
214 Research, Droevendaalsesteeg 3, Wageningen, P.O. Box 47, 6700 AA, The
215 Netherlands
- 216 68Escola de Negócios Tecnologia e Inovação, Centro Universitário do Pará, Belém,
217 PA, Brazil
- 218 69Universidade Federal do Pará, Rua Augusto Corrêa 01, Belém, PA, 66075-110,
219 Brazil
- 220 70Faculty of Natural Sciences, Department of Life Sciences, Imperial College
221 London, Silwood Park, South Kensington Campus, London, SW7 2AZ, UK
- 222 71Ecosistemas, Biodiversidad y Conservación de Especies, Universidad Estatal
223 Amazónica, Km. 2 1/2 vía a Tena (Paso Lateral), Puyo, Pastaza, Ecuador
- 224 72Museo de Historia Natural Noel Kempff Mercado, Universidad Autónoma Gabriel
225 Rene Moreno, Avenida Irala 565 Casilla Postal 2489, Santa Cruz, Santa Cruz,
226 Bolivia
- 227 73Universidade Federal de Minas Gerais, Instituto de Ciências Biológicas,
228 Departamento de Genética, Ecologia e Evolução, Av. Antônio Carlos, 6627
229 Pampulha, Belo Horizonte, MG, 31270-901, Brazil
- 230 74Department of Biology, University of Miami, Coral Gables, FL, 33146, USA
- 231 75Fairchild Tropical Botanic Garden, Coral Gables, FL, 33156, USA
- 232 76Instituto de Biociências - Dept. Ecologia, Universidade de Sao Paulo - USP, Rua do
233 Matão, Trav. 14, no. 321, Cidade Universitária, São Paulo, SP, 05508-090, Brazil
- 234 77Lancaster Environment Centre, Lancaster University, Lancaster, Lancashire, LA1
235 4YQ, UK
- 236 78Environmental Change Institute, University of Oxford, Oxford, Oxfordshire, OX1
237 3QY, UK
- 238 79Empresa Brasileira de Pesquisa Agropecuária, Embrapa Amapá, Rod. Juscelino
239 Kubitschek km 5, Macapá, Amapá, 68903-419, Brazil
- 240 80Grupo de Ecología y Conservación de Fauna y Flora Silvestre, Instituto Amazónico
241 de Investigaciones Imani, Universidad Nacional de Colombia sede Amazonia, Leticia,
242 Amazonas, Colombia
- 243 81Universidad Regional Amazónica IKIAM, Km 7 via Muyuna, Tena, Napo, Ecuador
- 244 82School of Environmental Sciences, University of East Anglia, Norwich, NR4 7TJ,
245 UK

- 246 83Dirección de Evaluación Forestal y de Fauna Silvestre, Av. Javier Praod Oeste 693,
247 Magdalena del Mar, Peru
248 84Escuela de Biología Herbario Alfredo Paredes, Universidad Central, Ap. Postal
249 17.01.2177, Quito, Pichincha, Ecuador
250 85Department of Biological Sciences, Humboldt State University, 1 Harpst Street,
251 Arcata, CA, 95521, USA
252 86Museu Universitário / Centro de Ciências Biológicas e da Natureza / Laboratório de
253 Botânica e Ecologia Vegetal, Universidade Federal do Acre, Rio Branco, AC, 69915-
254 559, Brazil
255 87Institute of Biological and Health Sciences, Federal University of Alagoas, Av.
256 Lourival Melo Mota, s/n, Tabuleiro do Martins, Maceio, AL, 57072-970, Brazil
257 88Iwokrama International Centre for Rain Forest Conservation and Development,
258 Georgetown, Guyana
259 89New York Botanical Garden, 2900 Southern Blvd, Bronx, New York, NY, 10458-
260 5126, USA
261 90School of Geosciences, University of Edinburgh, 201 Crew Building, King's
262 Buildings, Edinburgh, EH9 3JN, UK
263 91Tropical Diversity Section, Royal Botanic Garden Edinburgh, 20a Inverleith Row,
264 Edinburgh, Scotland, EH3 5LR, UK
265 92Servicios de Biodiversidad EIRL, Jr. Independencia 405, Iquitos, Loreto, 784, Peru
266 93Herbario Nacional de Bolivia, Universitario UMSA, Casilla 10077 Correo Central,
267 La Paz, La Paz, Bolivia
268 94Center for Conservation and Sustainable Development, Missouri Botanical Garden,
269 P.O. Box 299, St. Louis, MO, 63166-0299, USA
270 95Universidad Nacional de Jaén, Carretera Jaén San Ignacio Km 23, Jaén, Cajamarca,
271 06801, Peru
272 96Biology Department and Center for Energy, Environment and Sustainability, Wake
273 Forest University, 1834 Wake Forest Rd, Winston Salem, NC, 27106, USA
274 97Laboratoire Evolution et Diversité Biologique, CNRS and Université Paul Sabatier,
275 UMR 5174 EDB, Toulouse, 31000, France
276 98Andes to Amazon Biodiversity Program, Madre de Dios, Madre de Dios, Peru
277 99Department of Anthropology, University of Texas at Austin, SAC 5.150, 2201
278 Speedway Stop C3200, Austin, TX, 78712, USA
279 100Fundación Puerto Rastrojo, Cra 10 No. 24-76 Oficina 1201, Bogotá, DC,
280 Colombia
281 101Colegio de Ciencias Biológicas y Ambientales-COCIBA & Galapagos Institute
282 for the Arts and Sciences-GAIAS, Universidad San Francisco de Quito-USFQ, Quito,
283 Pichincha, Ecuador
284 102Department of Wildlife Ecology and Conservation, University of Florida, 110
285 Newins-Ziegler Hall, Gainesville, FL, 32611, USA
286 103Biosystematics group, Wageningen University, Droevendaalsesteeg 1,
287 Wageningen, 6708 PB, The Netherlands
288 104Fundación Estación de Biología, Cra 10 No. 24-76 Oficina 1201, Bogotá, DC,
289 Colombia
290 105Direction régionale de la Guyane, ONF, Cayenne, F-97300, French Guiana
291 106PROTERRA, Instituto de Investigaciones de la Amazonía Peruana (IIAP), Av. A.
292 Quiñones km 2,5, Iquitos, Loreto, 784, Peru
293 107ACEER Foundation, Jirón Cusco N° 370, Puerto Maldonado, Madre de Dios,
294 Peru

- 295 108Departement EV, Muséum national d'histoire naturelle de Paris, 16 rue Buffon,
296 Paris, 75005, France
- 297 109Amazon Conservation Team, Doekhieweg Oost #24, Paramaribo, Suriname
- 298 110Institut de Ciència i Tecnologia Ambientals, Universitat Autònoma de Barcelona,
299 08193 Bellaterra, Barcelona, Spain
- 300 111Environmental Change Institute, Oxford University Centre for the Environment,
301 Dyson Perrins Building, South Parks Road, Oxford, England, OX1 3QY, UK
- 302 112Instituto de Ciencias Naturales, Universidad Nacional de Colombia, Apartado
303 7945, Bogotá, DC, Colombia
- 304 113Instituto de Ciência Agrárias, Universidade Federal Rural da Amazônia, Av.
305 Presidente Tancredo Neves 2501, Belém, PA, 66.077-830, Brazil
- 306 114Escuela Profesional de Ingeniería Forestal, Universidad Nacional de San Antonio
307 Abad del Cusco, Jirón San Martín 451, Puerto Maldonado, Madre de Dios, Peru
- 308 115Universidad Autónoma del Beni José Ballivián, Campus Universitario Final, Av.
309 Ejercito, Riberalta, Beni, Bolivia
- 310 116Laboratory of Human Ecology, Instituto Venezolano de Investigaciones
311 Científicas - IVIC, Ado 20632, Caracas, DC, 1020A, Venezuela
- 312 117Cambridge University Botanic Garden, 1 Brookside., Cambridge, CB2 1JE, UK
- 313 118School of Agriculture and Food Sciences - ARC Centre of Excellence for
314 Environmental Decisions CEED, The University of Queensland, St. Lucia, QLD
315 4072, Australia
- 316 119University of Campinas, Plant Biology Department, Rua Monteiro Lobato, 255,
317 Cidade Universitária Zeferino Vaz, Barão Geraldo, Campinas, São Paulo, CEP
318 13083-862, Brazil
- 319 120Resource Ecology Group, Wageningen University & Research,
320 Droevendaalsesteeg 3a, Lumen, building number 100, Wageningen, Gelderland, 6708
321 PB, The Netherlands
- 322 121Laboratório de Ciências Ambientais, Universidade Estadual do Norte Fluminense,
323 Av. Alberto Lamego 2000, Campos dos Goyatacazes, RJ, 28013-620, Brazil
- 324 122Instituto de Investigaciones para el Desarrollo Forestal (INDEFOR), Universidad
325 de los Andes, Conjunto Forestal, 5101, Mérida, Mérida, Venezuela
- 326 123Departamento de Biología, Universidade Federal do Amazonas - UFAM –
327 Instituto de Ciências Biológicas – ICB1, Av General Rodrigo Octavio 6200, Manaus,
328 AM, 69080-900, Brazil
- 329 124GeoIS, El Día 369 y El Telégrafo, 3° Piso, Quito, Pichincha, Ecuador
- 330 125Department of Ecology and Evolutionary Biology, University of Michigan, Ann
331 Arbor, MI, 48109, USA
- 332 126University of Nottingham, University Park, Nottingham, NG7 2RD, UK
- 333 127School of Environmental and Forest Sciences, University of Washington, Seattle,
334 WA, 98195-2100, USA
- 335 128Environmental Science and Policy, Northern Arizona University, Flagstaff, AZ,
336 86011, USA
- 337 129Geography and the Environment, University of Texas at Austin, 305 E. 23rd
338 Street, CLA building, Austin, TX, 78712, USA
- 339 130Medio Ambiente, PLUSPRETOL, Iquitos, Loreto, Peru
- 340 131The Mauritius Herbarium, Agricultural Services, Ministry of Agro-Industry and
341 Food Security, Reduit, 80835, Mauritius
- 342 132Department of Bioscience, Aarhus University, Building 1540 Ny Munkegade,
343 Aarhus C, Aarhus, DK-8000, Denmark

- 344 133Living Earth Collaborative, Washington University in Saint Louis, St. Louis, MO,
345 63130, USA
- 346 134Escuela de Ciencias Forestales (ESFOR), Universidad Mayor de San Simon
347 (UMSS), Sacta, Cochabamba, Bolivia
- 348 135FOMABO, Manejo Forestal en las Tierras Tropicales de Bolivia, Sacta,
349 Cochabamba, Bolivia
- 350 136Tropenbos International, Lawickse Allee 11 PO Box 232, Wageningen, 6700 AE,
351 The Netherlands
- 352 137School of Anthropology and Conservation, University of Kent, Marlowe Building,
353 Canterbury, Kent, CT2 7NR, UK
- 354 138Herbario Nacional del Ecuador, Universidad Técnica del Norte, Quito, Pichincha,
355 Ecuador
- 356 139Instituto de Biodiversidade e Floresta, Universidade Federal do Oeste do Pará,
357 Rua Vera Paz, Campus Tapajós, Santarém, PA, 68015-110, Brazil
- 358 140Facultad de Biología, Universidad Nacional de la Amazonia Peruana, Pevas 5ta
359 cdra, Iquitos, Loreto, Peru
- 360 141Department of Biology, University of Missouri, St. Louis, MO, 63121, USA
- 361 142Department of Biogeochemical Integration, Max-Planck-Institute for
362 Biogeochemistry, P.O. Box 10 01 64, Jena, 07701, Germany
- 363 143Functional Biogeography, Max-Planck-Institute for Biogeochemistry, P.O. Box
364 10 01 64, Jena, 07701, Germany
- 365 144Department of Ecology and Evolutionary Biology, UCLA, 621 Charles E. Young
366 Drive South, Box 951606, Los Angeles, CA, 90095, USA
- 367 145Systems Ecology, Vrije Universiteit Amsterdam, De Boelelaan 1087, Amsterdam,
368 1081 HV, The Netherlands
- 369
- 370
- 371

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

386

387 **Key words:** maximum entropy, information theory, biodiversity, Amazon rainforest.

388

389 **Introduction**

390 Drivers of species distributions and their predictions have been a long-standing search
391 in ecology, with approaches varying from deterministic [1–6] to neutral [7–11] and
392 almost everything in between (e.g. near-neutral, continuum or emergent-
393 neutral: [12,13]). Most models are based on prior assumptions of processes that drive
394 community dynamics. The Maximum Entropy Formalism (hereafter called MEF),
395 however, makes no such a-priori assumptions in generating predictions of
396 distributions, including those of species abundances [14,15,24,16–23]. It is a useful
397 construct to infer processes driving community dynamics given the constraints
398 imposed by prior knowledge (e.g. functional traits or summed regional abundances),
399 as it does not include any bias introduced by potentially unjustified assumptions (29).
400 Quantifying the relative importance of these distinct constraints can thus provide
401 additional answers to understand the complexity of community dynamics (see
402 Supporting Materials SM: boxes S1-S3). This is especially so because, although many
403 different tests are available that link variation in taxon abundances to 1) trait variation,
404 2) taxon turnover between habitats or environments and 3) the distance decay of
405 similarities between samples, none quantify the importance of these relative to each
406 other. The MEF as applied here, however, is capable of and designed to do exactly
407 this by decomposing variation to separate information explained by each of these
408 aspects in a four-step model (Box S2) (25). Its application to an unprecedented large
409 tree inventory database on genus level taxonomy consisting of > 2,000 1-ha plots
410 distributed over Amazonia [25] and a genus trait database of 13 key functional traits
411 representing global axes of plant strategies [26] allows us to advance the study of
412 Amazonian tree community dynamics from a new cross-disciplinary perspective.

413

414

415 **Results**

416 Principles from information theory [14,15,27] can be used in an ecological setting to
417 predict the most likely abundance state for each taxon while simultaneously
418 maximizing entropy based on constraints. Maximization of entropy allows
419 quantifying the information yield for each constraint and therefor identifies which
420 constraints reduce entropy the most. Here we specifically use Shipley's mathematical
421 model to quantify probabilities, similar to earlier studies [28,29].

422

423 *Predictive power of the four-step model*

424 Using a uniform prior and both CWM and CWV as constraints accounted for 23% on
425 average of total deviance between observed and predicted relative abundances
426 (measured by R^2_{KL} values, see Box S2 equation 5). Filtered by forest type this was
427 36% for podzol forests, *várzea* 25%, *igapó* 23%, swamp forests 34%, 24% and 21%
428 for Guyana Shield and Pebas *terra firme* respectively and 20% for Brazilian Shield
429 *terra firme* forests (see Table S1 for detailed decomposition). Using observed
430 metacommunity relative abundances as prior regardless of CWM or CWV values
431 accounted on average 58% for the combined dataset with all forest types between 50
432 and 60%, except for the Guyana Shield *terra firme* with 63%. Including both trait
433 constraints and the metacommunity prior performed slightly better for the combined
434 dataset (average 62%), with a minimum of 56% for *igapó* forests and a maximum of
435 66% for the Guyana Shield *terra firme* forests. To compensate for spurious
436 relationships between regional abundances and local trait constraints, regardless of
437 selection, explanatory power was regarded relative to model bias yielding the pure
438 trait and metacommunity effects (Box S3, Fig. 2 and Table S1). This lowered the
439 proportion of information accounted for and yielded average pure metacommunity

440 effects of 43% for the overall dataset ranging between 30 and 48% for each forest
441 type separately with pure trait effects explaining only 5% of information for the
442 combined dataset on average with for each forest type between 3 and 8%. Although
443 the latter was lowered substantially, the explanatory power did appear to be strongly
444 dependent on forest type. Supplementary material provides additional results relating
445 to the predictive power of each model as well as the spatial gradient between the pure
446 trait and metacommunity effect ratios (Figs. S2-3).

447

448 *Direction and strength of selection of trait-based constraints*

449 Each trait showed significant differences in lambda when compared between forest
450 types (Fig. S1). Scatterplots of CWM trait values versus lambda show that, in general,
451 higher lambda values correspond with higher CWM trait values (Figure S7), although
452 the relationships are complex. Greater trait dissimilarity also correlated positively
453 with proportion of deviance attributable to pure trait effects (Pearson R of .23; Fig.
454 S9), supporting a trait-based selection. All traits except leaf nitrogen content also
455 showed reduction of variance accompanied a strong difference in community
456 weighted mean values associated with lambdas (either positive or negative), in line
457 with expectations of trait-based selection (Fig. S10).

458

459 *Effect of regional metacommunity prior*

460 There was a remarkable similar mean 21% decrease of the information explained
461 purely by the metacommunity prior for each forest type (Fig. 3). It should be noted
462 there is an obvious risk that when sampling size is increased, this also includes more
463 environmental heterogeneity as samples are coming from a variety of localities
464 potentially leading to changing composition. If this were the case, however, the

465 regional prior (q_i from Fig. S1 and Box S2) would also change, as taxa might be
466 abundant in some places but rare or absent in others. As the metacommunity effect is
467 the explained information that remains relative to any trait effects (i.e. information
468 unique to the neutral prior) and the pure trait effects are the explained information
469 remaining after correcting for pure metacommunity effects (Box S3) this effect should
470 then be accompanied by an increase in pure trait effect for each sample. This was not
471 observed, not even within the different forest types. Instead, the trait effect gradually
472 went up and then remained constant (Fig. S4).

473

474 **Discussion**

475 The MEF emerges from a well-founded theoretical and empirical body of ecology and
476 evolutionary biology, regarding natural selection, migration and population
477 dynamics [16,22,29,42,43]. From an ecological point of view it can be used to
478 quantify the relative association between directional or stabilizing selection for
479 functional traits versus the importance of relative regional abundance regardless of
480 these traits by imposing these as constraints. Our results show that pure trait effects,
481 on average, explained only 5% of the information when all forest types were taken
482 together whereas the pure metacommunity effect, however, explained almost ten
483 times more with an average value of 43%. Greater trait dissimilarity was positively
484 associated with higher pure trait effects, indicating trait-based selection, although the
485 assumed influence of dispersal regardless of these traits appeared to confer more
486 information explaining tree genus composition of the Amazon rainforest. The strength
487 and direction of selection indicated clear selective pressure for life history strategies
488 of either growth or protection, depending on forest type (see appendix S-A for a
489 detailed exploration of ecological interpretation).

490 Despite showing clear patterns in environmental selection and dispersal effects, there
491 was a large proportion of information left unexplained (44% on average). Potentially,
492 local demographic stochasticity could weaken any link between functional traits
493 measured and regional abundances of genera. This would, however, mean that almost
494 half of the information contained in relative abundances are the result of random
495 population dynamics and are not structurally governed. Alternatively, this could be
496 due to functional traits reflective of processes not taken into account in this study,
497 such as traits reflective of interactions between trophic levels. Another and at least
498 equally likely hypothesis for (local) unexplained information is that when scaling up,
499 the ratio of genus richness to total abundance decreases rather rapidly at first but
500 levels out as at some point relatively non-overlapping habitats are included in the
501 regional abundance distributions and more genera are included again due to the
502 different habitats. This would result in a change of the regional abundance distribution
503 (the prior) to which each local community is compared, resulting in higher local
504 unexplained information. Further study into these aspects could provide additional
505 insight, although as of yet the data necessary for these scales is lacking.

506 Although the initial explanatory power of the metacommunity prior differed between
507 forest types, the decay pattern was very similar. As the effects of either traits or the
508 metacommunity are measured in the goodness-of-fit predictions on local relative
509 abundances, this implies that at small spatial scales the surrounding regional
510 abundances provide better estimators than functional traits, while at larger spatial
511 scales this shifts to the traits. The ecological translation would be that on small spatial
512 scales, local communities share similar environmental conditions leaving dispersal
513 and drift acting in changing community composition, at least for genus level
514 taxonomy. As the potential regional pool is increased, more and more environmental

515 heterogeneity and non-overlapping regions are likely to be introduced. The more
516 gradual decline of *terra firme* forests can then arguably be attributed to these forests
517 having the largest relative surface area of Amazonia (even for the separate
518 subregions), potentially giving these forests an almost continuous metacommunity
519 without gaps, resulting in a more gradual transition from metacommunity to trait
520 relative importance. The fact the metacommunity effects do not change anymore after
521 certain distances would indicate the effect of dispersal potentially occurs over very
522 large distances. It should be noted that as these calculations are done at community
523 and genus level, they do not measure single dispersal events but rather the effect of
524 dispersal on community composition much deeper in time. In other words, this effect
525 suggests more than a dispersal event every now and then. Instead, it argues for
526 prolonged mixing of forests on large geographical and temporal scales, supported by
527 recent findings demonstrating a lack of geographical phylogenetic structure of
528 lineages for Amazonian tree genera [44].

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

539

540 **Methods**

541 *Empirical data*

542 The ATDN (ter Steege et al.) consists of over 2000 tree inventory plots distributed
543 over the Amazon basin and the Guiana Shield, collectively referred to as Amazonia.
544 Only those plots with trees ≥ 10 cm diameter at breast height were used, leaving 2011
545 plots with a mean of 558 individuals per plot identified to at least genus level. Most
546 plots used are 1 ha in size (1414) with 492 being smaller (minimum size of .1 ha) and
547 105 larger (maximum size of 80 ha). Genera have been standardized to the W3
548 Tropicos database (“Tropicos Missouri Botanical Garden”) using the Taxonomic
549 Name Resolution Service (TNRS [32]). After filtering based on above criteria and
550 solving nomenclature issues, 1,121,935 individuals belonging to over 828 genera
551 remained. Plots were distributed over seven abiotically different forest types: Podzol
552 forests (PZ), *Igapó* (IG, black water flood forests), *Várzea* (VA white water flood
553 forests), Swamp (SW) and *Terra firme* forests (TF) with subregions BS (Brazilian
554 Shield), GS (Guyana Shield) and PB (Pebas).

555

556 *Functional traits and trait imputation*

557 Constraints were formed by community weighted means (CWM) and variance of
558 functional traits (CWV), related to key ecological life history aspects on which natural
559 selection potentially operates (Table 1). According to principles of natural selection,
560 CWM values will be biased towards favourable trait values for that particular
561 environment in the case of directional selection, as taxa with these traits will be more
562 abundant due to environmental selection while stabilizing selection would decrease
563 CWV values [33]. For many traits it has been shown earlier that the interspecific
564 variability was larger than the intraspecific variability, allowing the use of data from

565 different sources to at least calculate a mean species trait value [34]. Genus trait
566 values were computed as genus-level means of species values if known within the
567 genus and considered constant for each genus. Genus level of taxonomy was used as
568 the available trait database had the most information on this taxonomic level.
569 Unknown values for traits were estimated by Multiple Imputation with Chained
570 Equations (MICE) by delta adjustment, subtracting a fixed amount (delta), with
571 sensitivity of this adjustment to the imputations of the observed versus imputed data
572 analysed using density plots (Fig. S8) and a linear regression model [35]. Procedure
573 was done using the *mice* package available on the R repository [36] under predictive
574 mean matching (*pmm* setting, 50 iterations). Results showed imputations were stable
575 and showed near identical patterns with each imputation scenario (see Figs S5-6 and
576 Table S2). After imputation, all trait values were transformed to Community
577 Weighted Means (CWM) of each trait (J) for each plot (K) (\bar{T}_{JK}) as $\bar{T}_{JK} =$
578 $\sum_{i=1}^S t_{ij} r a_{ik}$ with ra the relative abundance of the i^{th} genus in the k^{th} plot, following
579 earlier uses [37].
580

581 *MEF procedure predictions and ecological inference*

582 Figure 1 provides a schematic procedure overview, box S1 provides an overview of
583 important terms and Boxes S2-3 mathematical details. Initially, a maximally
584 uninformative prior is specified, where q_i (Box S1 equation 1) equals $1/S$ and trait
585 constraints are randomly permuted multiple times among genera to test whether
586 inclusion of specified constraints significantly changes derived probability
587 distributions (see also Roxburgh & Mokany, 2010). Subsequently, the same prior is
588 used but now observed trait CWM or CWV belonging to specific genera are used as
589 constraints. Third, observed regional abundances are used as prior with permuted
590 trait constraints and finally both observed regional abundances and observed trait
591 CWM/CWV are used as prior and constraints. *Maxent2* [28,29], an updated version
592 of the *maxent* function currently in the FD library of R [39] provided the
593 computational platform. Proportions of uncertainty explained by each model are given
594 by the Kullback-Leibler divergence R^2_{KL} , a generalization of the classic R^2 goodness
595 of fit [29]. Pure trait, pure metacommunity, joint metacommunity-trait and
596 unexplained effects are calculated as proportions of total biologically relevant
597 information (Box S1 and Box S2). Data was rarefied to smallest sample size (swamp
598 forests; 28) and calculations bootstrapped 25 times. Results indicated no significant
599 change compared to using all data, hence the total dataset was used for all analyses.

600

601 *Strength and direction of selection*

602 Predictions of genus relative abundances are computed as a function of traits reflected
603 in the CWM or CWV values and a series of constants (λ_{jk} : the Lagrange Multipliers).
604 Each multiplier quantifies the association between a unit of change for a particular
605 trait j and a proportional change in predicted relative abundance p_{ik} (the i^{th} genus in

606 the k^{th} community) considering all other traits are constant, formally described as:

607 $\frac{\partial p_{ik}}{\partial t_{ij}} = \lambda_{jk} p_{ik} (1 - p_{ik})$ (see appendix 1 from Sonnier, Navas, Fayolle, &

608 Shipley, 2012). Positive values indicate larger trait values associated with higher

609 abundances (positive selection), negative values indicate the opposite (negative

610 selection) with changes proportional to lambda. Values approximating zero indicate

611 no association between specific traits and relative abundances of species.

612 Decomposing λ_{jk} and comparing by means of a One-Way Analysis of Variance for

613 each trait separately between forest types allows studying both the strength and

614 direction of selection in different habitats. Note that this is done for the same

615 constraint between forest types, as lambda values for each constraint do not scale

616 linearly between different constraints.

617

618 *Estimation of metacommunity size*

619 Iteratively increasing the regional species pool taken into account as prior in

620 concentric circles of a fixed radius of 50 km allows estimating the spatial effect of

621 metacommunity size. The relationship between pure metacommunity effect and radius

622 of metacommunity size was fitted using a smoothing loess regression (function *loess*

623 and *predict*; R-package *stats* [41] with span set at 0.1). Fits subsequently were used

624 to predict values of metacommunity effect based on geographical distance to visualize

625 general patterns for each forest type. Exponential decay of pure metacommunity

626 effect was described using a self-start asymptotic regression function (*SSasymp*) of

627 the form $y(t) \sim y_f + (y_0 - y_f)e^{-\exp(\log(a))t}$ (*nls* from *stats*, R Core Team, 2016).

628

629

630

631

632 **References and notes:**

- 633 [1] D. Tilman, *Resource Competition between Plankton Algae: An Experimental*
634 *and Theoretical Approach*, Ecology **58**, 338 (1977).
- 635 [2] D. Tilman, *Resource Competition and Community Structure* (Princeton
636 university press, 1982).
- 637 [3] R. H. MacArthur, *Geographical Ecology: Patterns in the Distribution of*
638 *Species* (Princeton University Press, 1972).
- 639 [4] J. M. Chase and M. a Leibold, *Spatial Scale Dictates the Productivity-*
640 *Biodiversity Relationship.*, Nature **416**, 427 (2002).
- 641 [5] P. Chesson, *Mechanisms of Maintenance of Species Diversity*, Annu. Rev.
642 Ecol. Syst. **31**, 343 (2000).
- 643 [6] D. Tilman and S. W. Pacala, *The Maintenance of Species Richness in Plant*
644 *Communities*, Species Diversity in Ecological Communities.
- 645 [7] S. P. Hubbell, *The Unified Neutral Theory of Biodiversity and Biogeography*.
646 (Princeton Monographs in Population Biology. Princeton University Press,
647 Princeton, New Jersey, USA, 2001).
- 648 [8] I. Volkov, J. R. Banavar, S. P. Hubbell, and A. Maritan, *Neutral Theory and*
649 *Relative Species Abundance in Ecology.*, Nature **424**, 1035 (2003).
- 650 [9] J. Rosindell and S. J. Cornell, *Universal Scaling of Species-Abundance*
651 *Distributions across Multiple Scales*, Oikos **122**, 1101 (2012).
- 652 [10] G. Bell, *The Distribution of Abundance in Neutral Communities.*, Am. Nat.
653 **155**, 606 (2000).
- 654 [11] G. Bell, *The Co-Distribution of Species in Relation to the Neutral Theory of*
655 *Community Ecology*, Ecology.

656

- 657 [12] R. a Chisholm and S. W. Pacala, *Niche and Neutral Models Predict*
658 *Asymptotically Equivalent Species Abundance Distributions in High-Diversity*
659 *Ecological Communities.*, Proc. Natl. Acad. Sci. U. S. A. **107**, 15821 (2010).
- 660 [13] D. Tilman, *Niche Tradeoffs, Neutrality, and Community Structure: A*
661 *Stochastic Theory of Resource Competition, Invasion, and Community*
662 *Assembly.*, Proc. Natl. Acad. Sci. U. S. A. **101**, 10854 (2004).
- 663 [14] E. T. Jaynes, *Information Theory and Statistical Mechanics I*, Phys. Rev. **106**,
664 620 (1957).
- 665 [15] E. T. Jaynes, *Information Theory and Statistical Mechanics. II*, Phys. Rev. **108**,
666 171 (1957).
- 667 [16] B. Shipley, D. Vile, and E. Garnier, *From Plant Traits to Plant Communities:*
668 *A Statistical Mechanistic Approach to Biodiversity*, Science (80-.). **314**, 812
669 (2006).
- 670 [17] B. J. McGill, *A Renaissance in the Study of Abundance*, Solid State Phys. **314**,
671 770 (2006).
- 672 [18] S. Pueyo, F. He, and T. Zillio, *The Maximum Entropy Formalism and the*
673 *Idiosyncratic Theory of Biodiversity*, Ecol. Lett. **10**, 1017 (2007).
- 674 [19] R. C. Dewar and A. Porté, *Statistical Mechanics Unifies Different Ecological*
675 *Patterns*, J. Theor. Biol. **251**, 389 (2008).
- 676 [20] B. Shipley, *Inferential Permutation Tests for Maximum Entropy Models in*
677 *Ecology*, Ecology **91**, 2794 (2010).
- 678 [21] J. Bertram and R. C. Dewar, *Combining Mechanism and Drift in Community*
679 *Ecology: A Novel Statistical Mechanics Approach*, Theor. Ecol. **8**, 419 (2015).
- 680 [22] J. Harte, *Maximum Entropy and Ecology: A Theory of Abundance,*
681 *Distribution, and Energetics* (OUP Oxford, 2011).
- 682

- 683 [23] J. Harte and E. A. Newman, *Maximum Information Entropy: A Foundation for*
684 *Ecological Theory*, Trends Ecol. Evol. **29**, 384 (2014).
- 685 [24] J. P. O'Dwyer, A. Rominger, and X. Xiao, *Reinterpreting Maximum Entropy in*
686 *Ecology: A Null Hypothesis Constrained by Ecological Mechanism*, Ecol. Lett.
687 **20**, 832 (2017).
- 688 [25] H. ter Steege, N. C. A. Pitman, D. Sabatier, C. Baraloto, R. P. Salomão, J. E.
689 Guevara, O. L. Phillips, C. V. Castilho, W. E. Magnusson, J.-F. Molino, A.
690 Monteagudo, P. Núñez Vargas, J. C. Montero, T. R. Feldpausch, E. N. H.
691 Coronado, T. J. Killeen, B. Mostacedo, R. Vasquez, R. L. Assis, J. Terborgh, F.
692 Wittmann, A. Andrade, W. F. Laurance, S. G. W. Laurance, B. S. Marimon,
693 B.-H. Marimon, I. C. Guimarães Vieira, I. L. Amaral, R. Brienen, H.
694 Castellanos, D. Cárdenas López, J. F. Duivenvoorden, H. F. Mogollón, F. D. de
695 A. Matos, N. Dávila, R. García-Villacorta, P. R. Stevenson Diaz, F. Costa, T.
696 Emilio, C. Levis, J. Schiatti, P. Souza, A. Alonso, F. Dallmeier, A. J. D.
697 Montoya, M. T. Fernandez Piedade, A. Araujo-Murakami, L. Arroyo, R.
698 Gribel, P. V. A. Fine, C. A. Peres, M. Toledo, G. A. Aymard C, T. R. Baker, C.
699 Cerón, J. Engel, T. W. Henkel, P. Maas, P. Petronelli, J. Stropp, C. E. Zartman,
700 D. Daly, D. Neill, M. Silveira, M. R. Paredes, J. Chave, D. de A. Lima Filho, P.
701 M. Jørgensen, A. Fuentes, J. Schöngart, F. Cornejo Valverde, A. Di Fiore, E.
702 M. Jimenez, M. C. Peñuela Mora, J. F. Phillips, G. Rivas, T. R. van Andel, P.
703 von Hildebrand, B. Hoffman, E. L. Zent, Y. Malhi, A. Prieto, A. Rudas, A. R.
704 Ruschell, N. Silva, V. Vos, S. Zent, A. A. Oliveira, A. C. Schutz, T. Gonzales,
705 M. Trindade Nascimento, H. Ramirez-Angulo, R. Sierra, M. Tirado, M. N.
706 Umaña Medina, G. van der Heijden, C. I. A. Vela, E. Vilanova Torre, C.
707 Vriesendorp, O. Wang, K. R. Young, C. Baider, H. Balslev, C. Ferreira, I.

708 Mesones, A. Torres-Lezama, L. E. Urrego Giraldo, R. Zagt, M. N. Alexiades,
709 L. Hernandez, I. Huamantupa-Chuquimaco, W. Milliken, W. Palacios Cuenca,
710 D. Pauletto, E. Valderrama Sandoval, L. Valenzuela Gamarra, K. G. Dexter, K.
711 Feeley, G. Lopez-Gonzalez, M. R. Silman, S. P. Hubbell, F. He, R. Condit, L.
712 Borda-de-Agua, J. Kellner, H. Ter Steege, G. A. Black, T. H. Dobzhansky, C.
713 Pavan, J. M. Pires, T. Dobzhansky, G. A. Black, M. J. G. Hopkins, M. J.
714 Costello, R. M. May, N. E. Stork, P. Haripersaud, H. ter Steege, J.-J. de
715 Granville, H. Chevillotte, M. Hoff, D. P. Bebber, M. A. Carine, J. R. Wood, A.
716 H. Wortley, D. J. Harris, G. T. Prance, G. Davidse, J. Paige, T. D. Pennington,
717 N. K. Robson, R. W. Scotland, B. J. McGill, R. S. Etienne, J. S. Gray, D.
718 Alonso, M. J. Anderson, H. K. Benecha, M. Dornelas, B. J. Enquist, J. L.
719 Green, F. He, A. H. Hurlbert, A. E. Magurran, P. A. Marquet, B. A. Maurer, A.
720 Ostling, C. U. Soykan, K. I. Ugland, E. P. White, R. J. Warren, D. K. Skelly,
721 O. J. Schmitz, M. A. Bradford, N. C. A. Pitman, J. W. Terborgh, M. R. Silman,
722 P. N. V, D. A. Neill, C. E. Cerón, W. A. Palacios, M. Aulestia, N. C. A.
723 Pitman, M. R. Silman, J. W. Terborgh, F. D. Lozano, M. W. Schwartz, M. W.
724 Schwartz, D. Simberloff, J. E. Richardson, R. T. Pennington, T. D. Pennington,
725 P. M. Hollingsworth, T. L. Couvreur, F. Forest, W. J. Baker, S. Cavers, C. W.
726 Dick, D. H. Janzen, S. A. Mangan, S. A. Schnitzer, E. A. Herre, K. M. Mack,
727 M. C. Valencia, E. I. Sanchez, J. D. Bever, W. Balée, D. G. Campbell, C.
728 Levis, P. F. de Souza, J. Schiatti, T. Emilio, J. L. P. V. Pinto, C. R. Clement, F.
729 R. C. Costa, D. A. Posey, E. Montoya, V. Rull, N. D. Stansell, M. B. Abbott, S.
730 Nogué, B. W. Bird, W. A. Díaz, C. Gomez-Navarro, C. Jaramillo, F. Herrera,
731 S. L. Wing, R. Callejas, C. H. McMichael, D. R. Piperno, M. B. Bush, M. R.
732 Silman, A. R. Zimmerman, M. F. Raczka, L. C. Lobato, H. ter Steege, P. P.

733 Haripersaud, O. S. Bánki, F. Schieving, S. J. Phillips, R. P. Anderson, R. E.
734 Schapire, S. J. Phillips, M. Dudik, C. A. Quesada, J. Lloyd, L. O. Anderson, N.
735 M. Fyllas, M. Schwarz, C. I. Czimczik, B. Rollet, H. ter Steege, N. C. Pitman,
736 O. L. Phillips, J. Chave, D. Sabatier, A. Duque, J. F. Molino, M. F. Prévost, R.
737 Spichiger, H. Castellanos, P. von Hildebrand, R. Vásquez, P. M. Fearnside, D.
738 Mouillot, D. R. Bellwood, C. Baraloto, J. Chave, R. Galzin, M. Harmelin-
739 Vivien, M. Kulbicki, S. Lavergne, S. Lavorel, N. Mouquet, C. E. Paine, J.
740 Renaud, W. Thuiller, G. Lopez-Gonzalez, S. L. Lewis, M. Burkitt, O. L.
741 Phillips, P. J. M. Maas, L. Y. T. Westra, H. Rainer, A. Q. Lobão, R. H. J.
742 Erkens, P. V. A. Fine, D. C. Daly, G. V. Muñoz, I. Mesones, K. M. Cameron,
743 K. J. Feeley, M. R. Silman, M. Dufrene, P. Legendre, J. Bunge, M. Fitzpatrick,
744 J. Bunge, L. Woodard, D. Böhning, J. A. Foster, S. Connolly, H. K. Allen, J.-P.
745 Z. Wang, B. G. Lindsay, A. Chao, R. K. Colwell, C.-W. Lin, N. J. Gotelli, C.
746 X. Mao, R. K. Colwell, U. Brose, N. D. Martinez, R. J. Williams, I. J. Good, A.
747 Chao, S.-M. Lee, I. Rocchetti, J. Bunge, D. Bohning, R. A. Fisher, A. S.
748 Corbet, C. B. Williams, F. W. Preston, J.-P. Wang, P. V. A. Fine, and R. H.
749 Ree, *Hyperdominance in the Amazonian Tree Flora*, *Science* **342**, 1243092
750 (2013).
751 [26] S. Díaz, J. Kattge, J. H. C. Cornelissen, I. J. Wright, S. Lavorel, S. Dray, B.
752 Reu, M. Kleyer, C. Wirth, I. Colin Prentice, E. Garnier, G. Bönisch, M.
753 Westoby, H. Poorter, P. B. Reich, A. T. Moles, J. Dickie, A. N. Gillison, A. E.
754 Zanne, J. Chave, S. Joseph Wright, S. N. Sheremet Ev, H. Jactel, C. Baraloto,
755 B. Cerabolini, S. Pierce, B. Shipley, D. Kirkup, F. Casanoves, J. S. Joswig, A.
756 Günther, V. Falczuk, N. Rüger, M. D. Mahecha, and L. D. Gorné, *The Global*
757 *Spectrum of Plant Form and Function*, *Nature* **529**, 167 (2016).
758

- 759 [27] C. E. Shannon, *A Mathematical Theory of Communication*, Bell Syst. Tech. J.
760 **27**, 379 (1948).
- 761 [28] B. Shipley, *Measuring and Interpreting Trait-Based Selection versus Meta-*
762 *Community Effects during Local Community Assembly*, J. Veg. Sci. **25**, 55
763 (2014).
- 764 [29] B. Shipley, C. E. T. Paine, and C. Baraloto, *Quantifying the Importance of*
765 *Local Niche-Based and Stochastic Processes to Tropical Tree Community*
766 *Assembly*, Ecology **93**, 760 (2012).
- 767 [30] H. ter Steege, A. Monteagudo, A. F. Sampaio, A. Ruschel, A. Prieto, A.
768 Quaresma, A. Rudas, A. Vicentini, A. Araujo-Murakami, A. A. de Oliveira, A.
769 Alonso, A. Fuentes, A. Lopes, A. Duque, A. S. de Andrade, A. Cano, A. G.
770 Manzatto, A. Di Fiore, A. Torres-Lezama, B. S. Marimon, B. H. M. Jr, B. W.
771 Albuquerque, B. Laurance, B. Magnusson, B. Mostacedo, B. Villa, B.
772 Hoffman, B. Cintra, B. G. Luize, C. Cerón, C. Peres, C. Castilho, C. Levis, C.
773 Mendoza, C. Vela, C. E. Zartman, C. Baraloto, C. Ferreira, C. Baider, C.
774 Vriesendorp, D. Cárdenas, D. Praia, D. Sabatier, D. Villarroel, D. Pauletto, D.
775 D. do Amaral, D. Neill, D. Correa, D. Reyes, D. de A. L. Filho, D. Rodrigues,
776 D. Daly, E. M. Barbosa, E. M. Venticinque, E. Zent, E. M. Jimenez, E.
777 Valderrama, E. de S. Farias, E. V. Torre, E. Berenguer, E. Honorio, E.
778 Almeida, E. M. M. de L. Novo, F. A. Carvalho, F. C. de Souza, F. C. Valverde,
779 F. Costa, F. Wittmann, F. D. de A. Matos, F. Dallmeier, F. R. Arevalo, G.
780 Damasco, G. Lopez-Gonzalez, G. van der Heijden, G. P. G. Gonzales, G.
781 Aymard, G. Rivas, H. Beltrán, H. Balslev, H. E. M. Nascimento, H.
782 Castellanos, H. P. D. Doza, H. Ramirez-Angulo, H. Mogollón, I. Amaral, I.
783 Vieira, I. P. de A. Miranda, I. Huamantupa-Chuquimaco, I. Mesones, J. Costa,

784 J.-F. Molino, J.-L. Guillaumet, J. Chave, J. Comiskey, J. Ricardo, J. Schöngart,
785 J. Terborgh, J. Ferreira, J. Lloyd, J. Duivenvoorden, J. Barlow, J. F. Ramos, J.
786 J. de Toledo, J. L. Camargo, J. L. M. Pena, J. R. da S. Guimarães, J. C. Licona,
787 J. C. Montero, J. D. C. Revilla, J. E. Guevara, J. F. Phillips, J. Schietti, J.
788 Stropp, J. Engel, K. Garcia, K. R. Casula, K. Feeley, K. R. Young, K. Dexter,
789 L. Silva, L. Demarchi, L. Ferreira, L. E. Urrego, L. F. M. Pinto, L. Hernandez,
790 L. Barbosa, L. V. Gamarra, L. Casas, L. C. de M. Bonates, L. de S. Coelho, L.
791 Arroyo, M. Rocha, M. Piedade, M. Ahuite, M. C. Guedes, M. de J. V. Carim,
792 M. Medeiros, M. P. Pansonato, M. Simon, M. T. Nascimento, M. dos Santos,
793 M. R. Paredes, M. Silveira, M. C. P. Mora, M. N. U. Medina, M. P. Martins,
794 M. V. Irumé, M. Alexiades, M. Silman, M. Aulestia, M. Tirado, N. Dávila, N.
795 Targhetta, N. Silva, N. F. C. Reis, N. C. Arboleda, N. Pitman, O. Bánki, O.
796 Phillips, O. Wang, P. Stevenson, P. Petronelli, P. von Hildebrand, P. Fine, P.
797 Maas, P. N. Vargas, P. M. Jørgensen, P. Souza, R. L. de Assis, R. Salomão, R.
798 Thomas-Caesar, R. R. Hilário, R. Linares-Palomino, R. Z. Gómez, R. Foster,
799 R. Zagt, R. Vasquez, R. Sierra, R. Brienen, R. Gribel, R. García, S. Cárdenas,
800 S. Mori, S. Zent, S. Pansini, S. Laurance, T. Feldpausch, T. Henkel, T. Emilio,
801 T. Gonzales, T. S. F. Silva, T. Baker, T. Killeen, T. van Andel, T. Gardner, T.
802 Pennington, V. Vos, W. Palacios, W. Campelo, W. Farfan, W. Milliken, W.
803 Nauray, Y. Malhi, and Y. Feitosa, *Http://Atdn.Myspecies.Info/*, (unpublished).
804 [31] *Tropicos Missouri Botanical Garden*, (unpublished).

805

- 806 [32] B. Boyle, N. Hopkins, Z. Lu, J. A. Raygoza Garay, D. Mozzherin, T. Rees, N.
807 Matasci, M. L. Narro, W. H. Piel, S. J. McKay, S. Lowry, C. Freeland, R. K.
808 Peet, and B. J. Enquist, *The Taxonomic Name Resolution Service: An Online*
809 *Tool for Automated Standardization of Plant Names.*, BMC Bioinformatics **14**,
810 16 (2013).
- 811 [33] J. Loranger, F. Munoz, B. Shipley, and C. Violle, *What Makes Trait–*
812 *Abundance Relationships When Both Environmental Filtering and Stochastic*
813 *Neutral Dynamics Are at Play?*, Oikos **127**, 1735 (2018).
- 814 [34] E. Kazakou, C. Violle, C. Roumet, M. L. Navas, D. Vile, J. Kattge, and E.
815 Garnier, *Are Trait-Based Species Rankings Consistent across Data Sets and*
816 *Spatial Scales?*, J. Veg. Sci. **25**, 235 (2014).
- 817 [35] S. Van Buuren, *Flexible Imputation of Missing Data* (Chapman and Hall/CRC,
818 2018).
- 819 [36] S. van Buuren and K. Groothuis-Oudshoorn, *Mice: Multivariate Imputation by*
820 *Chained Equations in R*, J. Stat. Softw. **45**, 1 (2011).
- 821 [37] B. Shipley, D. C. Laughlin, G. Sonnier, and R. Otfinowski, *A Strong Test of a*
822 *Maximum Entropy Model of Trait-Based Community Assembly*, Ecology **92**,
823 507 (2011).
- 824 [38] S. H. Roxburgh and K. Mokany, *On Testing Predictions of Species Relative*
825 *Abundance from Maximum Entropy Optimisation*, Oikos **119**, 583 (2010).
- 826 [39] E. Laliberté and B. Shipley, *FD: Measuring Functional Diversity from*
827 *Multiple Traits, and Other Tools for Functional Ecology. R Package Version*
828 *1.0-11*, URL [Http//CRAN.R-Project. Org/Package= FD](http://CRAN.R-project.org/package=FD) [Accessed 9 January
829 2012] (2011).

830

- 831 [40] G. Sonnier, M. L. Navas, A. Fayolle, and B. Shipley, *Quantifying Trait*
832 *Selection Driving Community Assembly: A Test in Herbaceous Plant*
833 *Communities under Contrasted Land Use Regimes*, *Oikos* **121**, 1103 (2012).
- 834 [41] R Core Team, *R: A Language and Environment for Statistical Computing*.
- 835 [42] B. Shipley, *From Plant Traits to Vegetation Structure. Chance and Selection in*
836 *the Assembly of Ecological Communities* (Cambridge University Press, 2010).
- 837 [43] M. Favretti, *Remarks on the Maximum Entropy Principle with Application to*
838 *the Maximum Entropy Theory of Ecology*, *Entropy* **20**, (2018).
- 839 [44] K. G. Dexter, M. Lavin, B. M. Torke, A. D. Twyford, T. A. Kursar, P. D.
840 Coley, C. Drake, R. Hollands, and R. T. Pennington, *Dispersal Assembly of*
841 *Rain Forest Tree Communities across the Amazon Basin*, *Proc. Natl. Acad. Sci.*
842 **114**, 2645 (2017).
- 843 [45] E. Weiher, A. van der Werf, K. Thompson, M. Roderick, E. Garnier, and O.
844 Eriksson, *Challenging Theophrastus: A Common Core List of Plant Traits for*
845 *Functional Ecology*, *J. Veg. Sci.* **10**, 609 (1999).
- 846 [46] K. G. Raghothama, *Phosphate Acquisition*, *Annu. Rev. Plant Physiol. Plant*
847 *Mol. Biol.* **50**, 665 (1999).
- 848 [47] L. Poorter, M. Van De Plassche, S. Willems, and R. G. A. Boot, *Leaf Traits*
849 *and Herbivory Rates of Tropical Tree Species Differing in Successional Status*,
850 *Plant Biol.* **6**, 746 (2004).
- 851 [48] A. A. Agrawal, *Macroevolution of Plant Defense Strategies*, *Trends Ecol. Evol.*
852 **22**, 103 (2007).
- 853 [49] J. I. Sprent and R. Parsons, *Nitrogen Fixation in Legume and Non-Legume*
854 *Trees*, *F. Crop. Res.* **65**, 183 (2000).

855

- 856 [50] F. Shah, C. Nicolas, J. Bentzer, M. Ellstrom, M. Smits, F. Rineau, B. Canback,
857 D. Floudas, R. Carleer, G. Lackner, J. Braesel, D. Hoffmeister, B. Henrissat, T.
858 Johansson, D. S. Hibbett, F. Martin, P. Persson, and A. Tunlid,
859 *Ectomycorrhizal Fungi Decompose Soil Organic Matter Using Oxidative*
860 *Mechanisms Adapted from Saprotrophic Ancestors*, *New Phytol.* **1705** (2017).
- 861 [51] J. V. Colpaert, J. H. L. Wevers, E. Krznaric, and K. Adriaensen, *How Metal-*
862 *Tolerant Ecotypes of Ectomycorrhizal Fungi Protect Plants from Heavy Metal*
863 *Pollution*, *Ann. For. Sci.* **68**, 17 (2011).
- 864 [52] C. Poschenrieder, B. Gunsé, I. Corrales, and J. Barceló, *A Glance into*
865 *Aluminum Toxicity and Resistance in Plants*, *Sci. Total Environ.* **400**, 356
866 (2008).
- 867 [53] H. F. Howe and J. Smallwood, *Ecology of Seed Dispersal Ecology of Seed*
868 *Dispersal*, **13**, 201 (2013).
- 869 [54] C. Baraloto, C. E. T. Paine, L. Poorter, J. Beauchene, D. Bonal, A. M.
870 Domenach, B. Hérault, S. Patiño, J. C. Roggy, and J. Chave, *Decoupled Leaf*
871 *and Stem Economics in Rain Forest Trees*, *Ecol. Lett.* **13**, 1338 (2010).
- 872 [55] C. Fortunel, P. V. A. Fine, and C. Baraloto, *Leaf, Stem and Root Tissue*
873 *Strategies across 758 Neotropical Tree Species*, *Funct. Ecol.* **26**, 1153 (2012).
- 874 [56] N. Fyllas and S. Patino, *Basin-Wide Variations in Foliar Properties of*
875 *Amazonian Forest: Phylogeny, Soils and Climate*, *Biogeosciences* 2677 (2009).
- 876 [57] N. J. B. Kraft, R. Valencia, and D. D. Ackerly, *Functional Traits and Niche-*
877 *Based Tree Community Assembly in an Amazonian Forest.*, *Science* **322**, 580
878 (2008).
- 879 [58] F. M. J. Ohler, *Phytomass and Mineral Content in Untouched Forest*, (1980).
- 880

- 881 [59] C. E. T. Paine, C. Baraloto, J. Chave, and B. Hérault, *Functional Traits of*
882 *Individual Trees Reveal Ecological Constraints on Community Assembly in*
883 *Tropical Rain Forests*, *Oikos* **120**, 720 (2011).
- 884 [60] T. L. Pons, K. Perreijn, C. Van Kessel, and M. J. A. Werger, *Symbiotic*
885 *Nitrogen Fixation in a Tropical Rainforest: ¹⁵N Natural Abundance*
886 *Measurements Supported by Experimental Isotopic Enrichment*, *New Phytol.*
887 **173**, 154 (2007).
- 888 [61] J. Thompson, J. Proctor, V. Viana, W. Milliken, J. A. Ratter, and D. A. Scott,
889 *Ecological Studies on a Lowland Evergreen Rain Forest on Maraca Island,*
890 *Roraima, Brazil. I. Physical Environment, Forest Structure and Leaf*
891 *Chemistry*, *J. Ecol.* 689 (1992).
- 892 [62] J. Kattge, S. Díaz, S. Lavorel, I. C. Prentice, P. Leadley, G. Bönisch, E.
893 Garnier, M. Westoby, P. B. Reich, I. J. Wright, J. H. C. Cornelissen, C. Violle,
894 S. P. Harrison, P. M. Van Bodegom, M. Reichstein, B. J. Enquist, N. A.
895 Soudzilovskaia, D. D. Ackerly, M. Anand, O. Atkin, M. Bahn, T. R. Baker, D.
896 Baldocchi, R. Bekker, C. C. Blanco, B. Blonder, W. J. Bond, R. Bradstock, D.
897 E. Bunker, F. Casanoves, J. Cavender-Bares, J. Q. Chambers, F. S. Chapin, J.
898 Chave, D. Coomes, W. K. Cornwell, J. M. Craine, B. H. Dobrin, L. Duarte, W.
899 Durka, J. Elser, G. Esser, M. Estiarte, W. F. Fagan, J. Fang, F. Fernández-
900 Méndez, A. Fidelis, B. Finegan, O. Flores, H. Ford, D. Frank, G. T. Freschet,
901 N. M. Fyllas, R. V. Gallagher, W. A. Green, A. G. Gutierrez, T. Hickler, S. I.
902 Higgins, J. G. Hodgson, A. Jalili, S. Jansen, C. A. Joly, A. J. Kerkhoff, D.
903 Kirkup, K. Kitajima, M. Kleyer, S. Klotz, J. M. H. Knops, K. Kramer, I. Kühn,
904 H. Kurokawa, D. Laughlin, T. D. Lee, M. Leishman, F. Lens, T. Lenz, S. L.
905 Lewis, J. Lloyd, J. Llusià, F. Louault, S. Ma, M. D. Mahecha, P. Manning, T.

906 Massad, B. E. Medlyn, J. Messier, A. T. Moles, S. C. Müller, K. Nadrowski, S.
907 Naeem, Ü. Niinemets, S. Nöllert, A. Nüske, R. Ogaya, J. Oleksyn, V. G.
908 Onipchenko, Y. Onoda, J. Ordoñez, G. Overbeck, W. A. Ozinga, S. Patiño, S.
909 Paula, J. G. Pausas, J. Peñuelas, O. L. Phillips, V. Pillar, H. Poorter, L. Poorter,
910 P. Poschlod, A. Prinzing, R. Proulx, A. Rammig, S. Reinsch, B. Reu, L. Sack,
911 B. Salgado-Negret, J. Sardans, S. Shiodera, B. Shipley, A. Siefert, E. Sosinski,
912 J. F. Soussana, E. Swaine, N. Swenson, K. Thompson, P. Thornton, M.
913 Waldram, E. Weiher, M. White, S. White, S. J. Wright, B. Yguel, S. Zaehle, A.
914 E. Zanne, and C. Wirth, *TRY - a Global Database of Plant Traits*, *Glob. Chang.*
915 *Biol.* **17**, 2905 (2011).

916

917 **Author Contributions:**

918 E.T. Pos and H. ter Steege designed the experiment. E.T.Pos wrote R-scripts,
919 analysed the results and took the lead in writing the manuscript, H. ter Steege
920 supervised the writing and provided regular feedback both for the manuscript and the
921 interpretation of the results. All other authors provided feedback on the manuscript
922 and provided their data from the Amazon Tree Diversity Network or trait data.
923 Authors E.T. Pos to L.V. Gamarra provided tree inventory data, authors G. Boenisch,
924 J. Kattge, N. Kraft, A. Levesley, K. Melgaço, G. Pickavance, L. Poorter provided data
925 on functional traits, C. Baraloto, J. Lloyd, A. A. Oliveira and H. ter Steege provided
926 both tree inventory and functional trait data. **Competing interests:** Authors declare
927 no competing interests. **Data and materials availability:** R scripts are available on
928 the github repository of E.T. Pos (EdwinTPos). The data that support the findings of
929 this study are available from The Amazon Tree Diversity Network (ATDN) upon
930 reasonable request.

931 **Acknowledgements:**

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

945

946 **Funding:** HtS and RS were supported by grant 407232/2013-3 - PVE -
947 MEC/MCTI/CAPES/CNPq/FAPs; CB was supported by grant FAPESP 95/3058-0 -
948 CRS 068/96 WWF Brasil - The Body Shop; DS, JFM, JE, PP and JC benefited from
949 an “Investissement d’Avenir” grant managed by the Agence Nationale de la
950 Recherche (CEBA: ANR-10-LABX-25-01); HLQ/MAP/JLLM received financial
951 supported by MCT/CNPq/CT-INFRA/GEOMA #550373/2010-1 and # 457515/2012-
952 0, and JLLM were supported by grant CAPES/PDSE # 88881.135761/2016-01 and
953 CAPES/Fapespa #1530801; The Brazilian National Research Council (CNPq)
954 provided a productivity grant to EMV (Grant 308040/2017-1); Floristic identification

955 in plots in the RAINFOR forest monitoring network have been supported by the
956 Natural Environment Research Council (grants NE/B503384/1, NE/ D01025X/1,
957 NE/I02982X/1, NE/F005806/1, NE/D005590/1 and NE/I028122/1) and the Gordon
958 and Betty Moore Foundation; B.M.F. is funded by FAPESP grant 2016/25086-3.
959 BSM, BHMJ and OLP were supported by grants CNPq/CAPES/FAPS/BC-Newton
960 Fund #441244/2016-5 and FAPEMAT/0589267/2016; TWH was funded by National
961 Science Foundation grant DEB-1556338.

TABLE 1

Functional trait	Units	Mean	SD	Est %	Associated challenge
Wood density (<i>WD</i>)	g/cm ³	0.63	0.17	30	Longevity [45]
Seed Mass Class (<i>SMC</i>)	categorical (1-8)	4.3	1.4	31	Dispersal, Fecundity, Establishment [45]
Specific Leaf Area (<i>SLA</i>)	mm ² /mg	15	5.9	41	Establishment, Plasticity, Disturbance [45]
Leaf nitrogen content (<i>N</i>)	mg/g	22.3	7.30	41	Photosynthetic capacity [45]
Leaf phosphorus content (<i>P</i>)	mg/g	1	0.77	50	Limited available P for metabolism [46]
Leaf carbon content (<i>C</i>)	mg/g	468	38.1	54	Herbivore resistance (C:N) [47]
Latex	1=no, 2=yes	1.2	0.43	46	Herbivore resistance [48]
Resin	1=no, 2=yes	1.1	0.35	58	Herbivore resistance [48]
Root Nodules (<i>Nodules</i>)	1=no, 2=yes	1.1	0.28	0	Nitrogen fixation [49]
Ectomycorrhiza (<i>EctoMyco</i>)	1=no, 2=yes	1.01	0.11	0	Organic N fixation [50], heavy metal pollution [51]
Aluminum accumulation (<i>AlAcc</i>)	1=no, 2=yes	1.1	0.21	3	Heavy metal pollution [52]
Fleshy Fruits (<i>Fleshy</i>)	1=no, 2=yes	1.6	0.50	7	Dispersal (<i>specificity</i>) [53]
Winged seeds (<i>Wings</i>)	1=no, 2=yes	1.2	0.42	39	Dispersal (<i>limitation</i>) [53]

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: Chris Baraloto (TRY), Adalardo de Oliveira (unpublished data), L. Poorter (unpublished data), J. Lloyd (TRY), Van der Sande and Mazzei (unpublished data), Van der Sande and Poorter (unpublished data), [54–62]

FIGURE 1

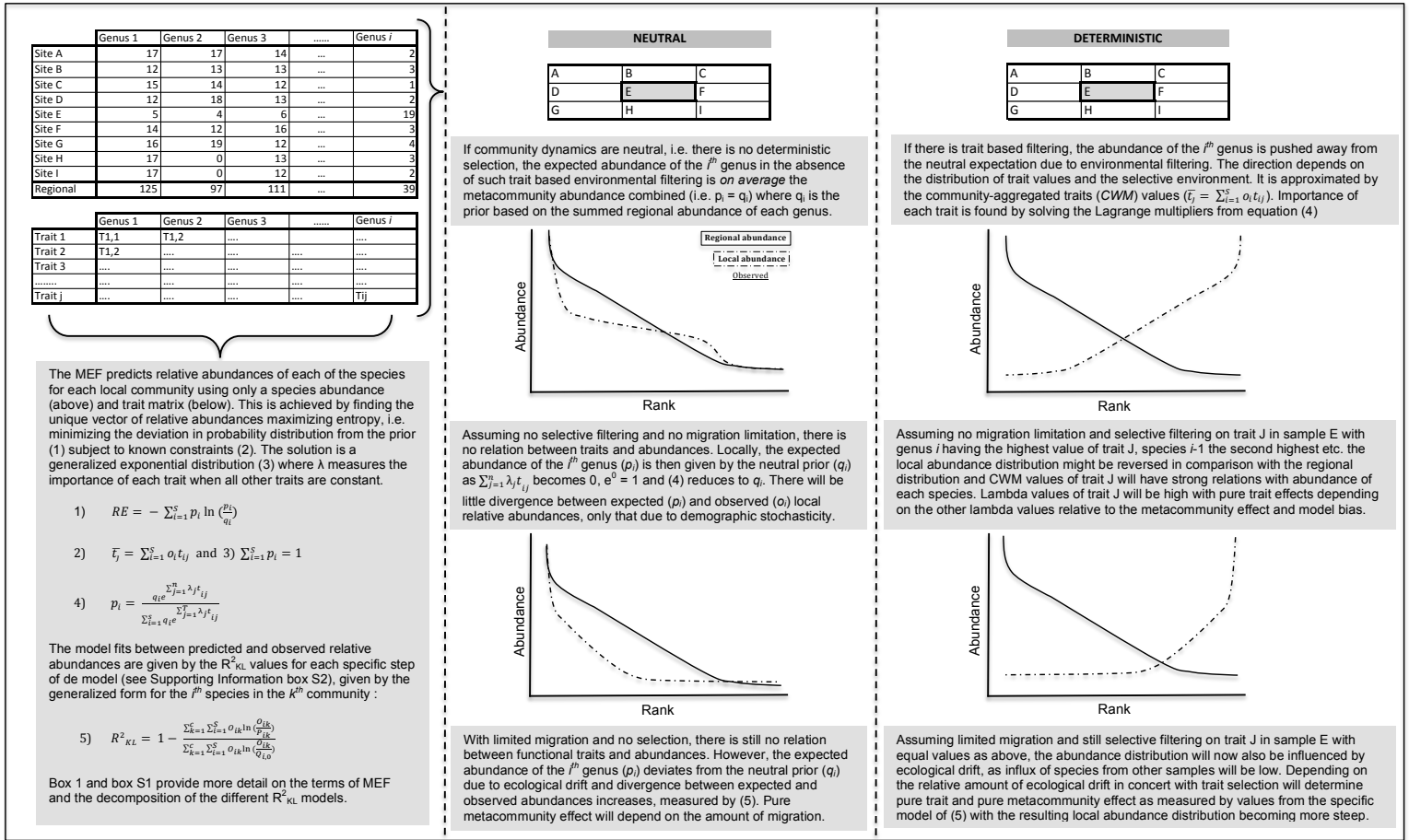


Fig. 1. Schematic depiction of the MEF procedure. Left panel shows a genus abundances per site and a functional trait matrix per genus, bottom half outlines calculations. Middle and right panel show different scenarios of neutral and deterministic dynamics under infinite or limited migration. Dashed and solid lines indicate local and regional abundance distributions respectively.

FIGURE 2

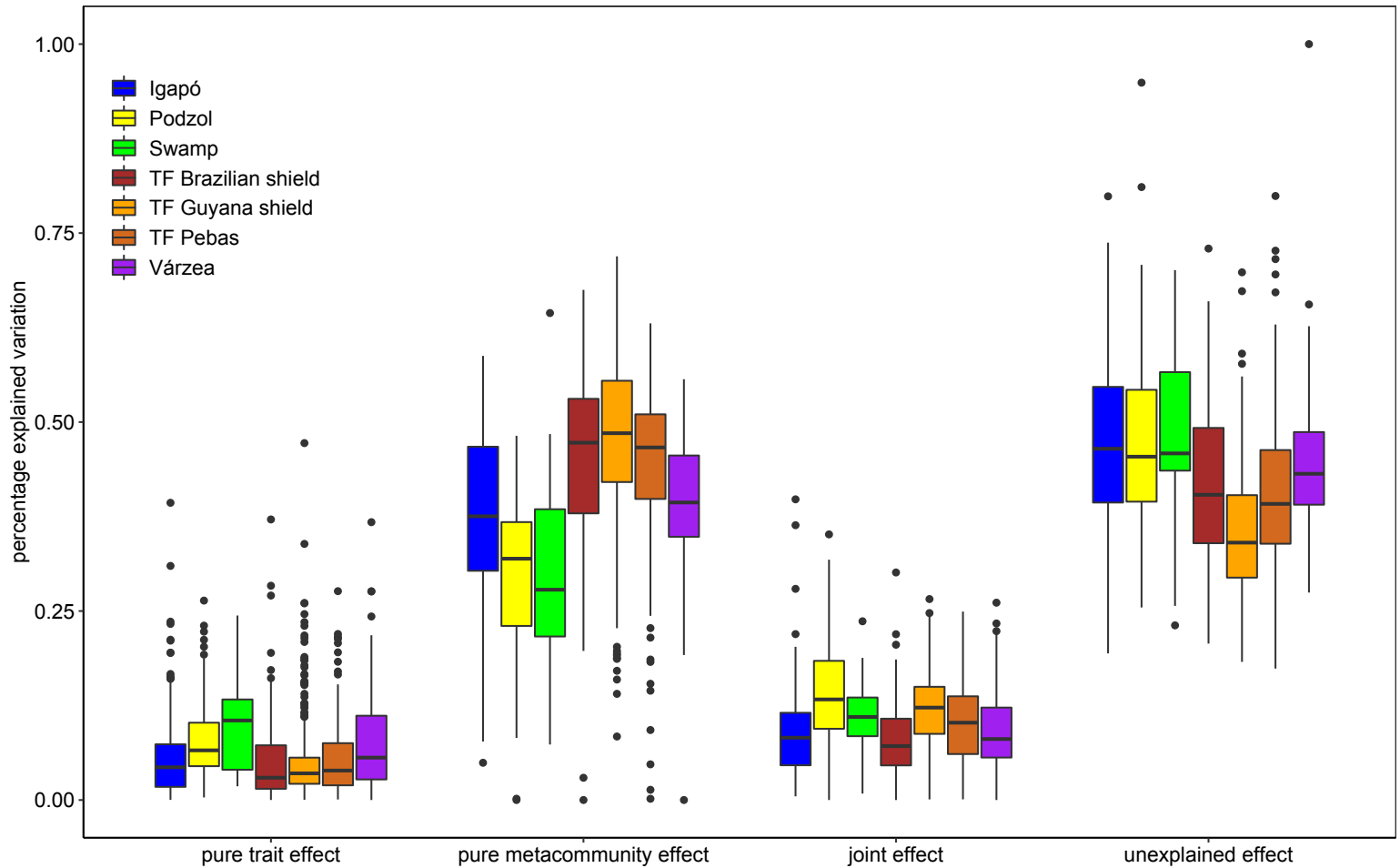


Fig. 2. Visual representation of pure trait, pure metacommunity, hybrid model and the remaining unexplained information for each separate forest type.

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 $1.5 * IQR$ from hinge. Points beyond this range are plotted individually.

FIGURE 3

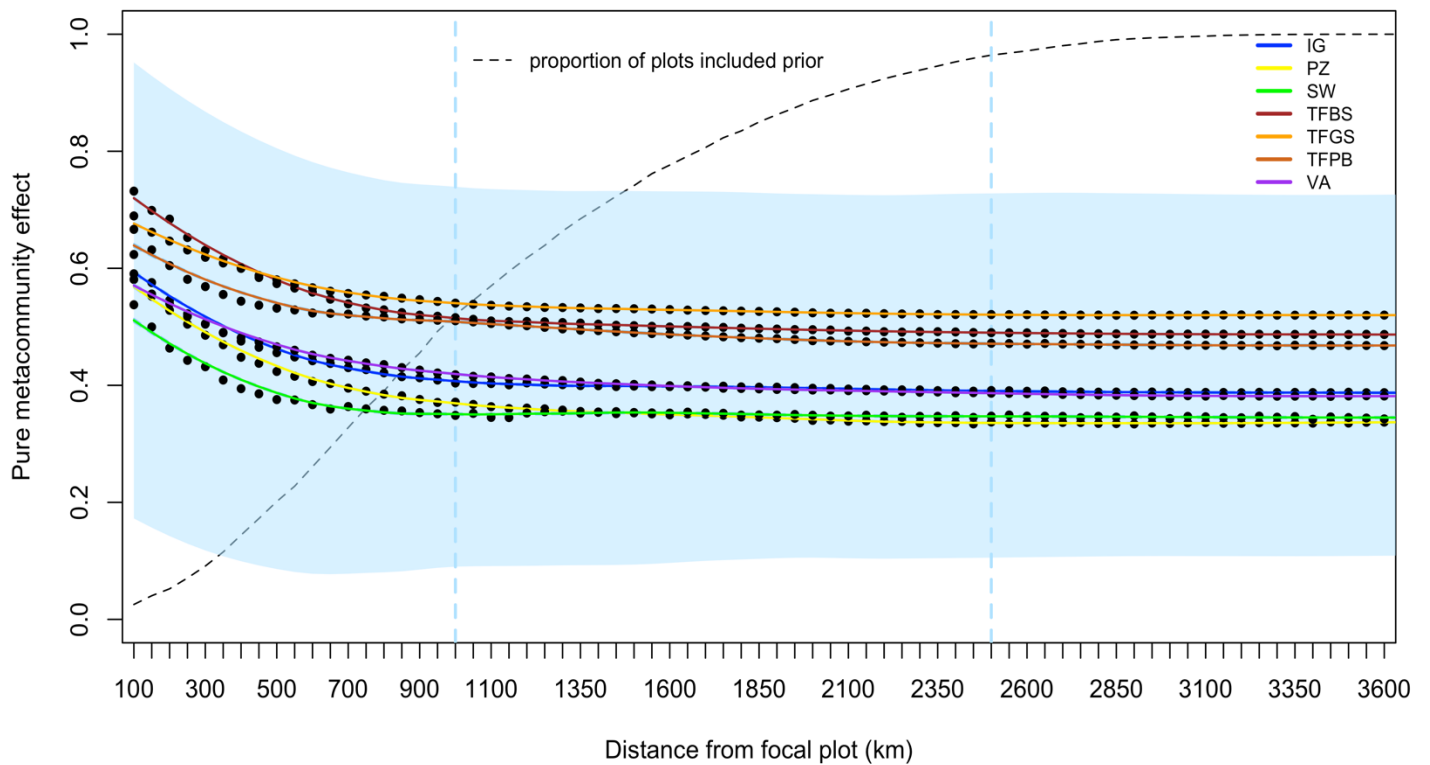


Fig. 3. Distance decay of pure metacommunity effect. X-axis represents radius of metacommunity prior. Dashed line indicates mean number of plots for that distance included as metacommunity prior. Y-axis represents information unique to metacommunity prior taken relative to model bias. Solid lines indicate predictions from loess regression based on all points with different colours indicating forest types with abbreviations as in main text. Blue shading reflects minimum and maximum loess regression predicted values.

BOX S1

Entities:

The basic unit of the MEF model can exist in different states. If the system under study is a collection of genera existing at a site, then each entity is a single genus.

States:

Classification of different ways any entity can exist. In the same collection of taxa, states of each entity (i.e. genus) are their specific abundance at that site. Microstates are the exact arrangement in time and space for the states of the entities in the system. Macrostates are the description of entities among the possible states in the system under study without regard to the spatial or temporal arrangement of these entities. I.e. observing a relative abundance distribution, but not the actual dispersal and germination of individuals.

Traits, attributes or properties:

Each entity possesses measurable properties whose values will probably differ between states. For example, genera differ in average wood density, seed mass, height etcetera.

Maximally uninformative prior:

All the information concerning states before constraints are introduced. Called maximally uninformative as preferably all empirical information is introduced in the form of constraints as to have the maximal gain of information regarding the different traits.

Prior distribution:

Prior distribution of expected states for the entities which can be incorporated as a constraint in addition to the traits, being either the observed relative abundance of each entity in the summed sample (i.e. the metacommunity) or a maximally uninformed (uniform) distribution. The former would be a neutral prior (expected local abundance is equal to the abundance in the larger metacommunity).

Community-weighted means:

The average trait value (i.e. measurable property such as wood density) of entities (such as genera) weighted by the relative abundance of each entity at a specific site

Box S1. Different ingredients necessary for analyses using MEF. Definitions of the most important terms used in the MEF analyses and throughout the main text to provide the necessary framework of understanding.

BOX S2

<p>The Maximum Entropy Formalism works on the basis of a conceptual model called the CATS (<i>Community Assembly by Trait Selection</i>) and makes use of three inputs:</p> <p>i) A trait matrix containing the measured functional traits of each of the S total genera in the total regional pool, these can be of either discrete or continuous form.</p> <p>ii) A vector of n community weighted trait values, estimating the average trait value over all individuals in the local community for each of the traits</p> <p>iii) A prior probability distribution specifying the regional abundance distribution, quantifying potential contributions of the regional pool of recruits to the structure of local communities.</p> <p>Using these three sources of information, the model predicts relative abundances (p_i) in the form of Bayesian probabilities for each genus in each local community without assuming any a priori relations or processes. This is achieved by finding the vector of relative abundances maximizing entropy:</p> $1) RE = - \sum_{i=1}^S p_i \ln \left(\frac{p_i}{q_i} \right)$ <p>with q_i the regional species pool abundance of species i and RE (Relative Entropy) subject to the known constraints for j traits and i species:</p> $2) \bar{t}_j = \sum_{i=1}^S o_i t_{ij} \text{ and } 3) \sum_{i=1}^S p_i = 1$ <p>The solution is a generalized exponential distribution where the λ values measure the importance of each trait when all other traits are constant:</p> $4) p_i = \frac{q_i e^{\sum_{j=1}^n \lambda_j t_{ij}}}{\sum_{i=1}^S q_i e^{\sum_{j=1}^n \lambda_j t_{ij}}}$ <p>Note that when all λ values are zero, i.e. there is no trait based selection, $p_i = q_i$</p>	<p>The final step is to measure the proportion of total deviance accounted for between observed and predicted relative abundances for each of the four-step solution. These are the R^2_{KL} values, a generalization of the classic R^2 index of maximum likelihood estimation using the Kullback-Leibler index [16]:</p> <p>i) $\bar{R}^2_{KL}(\mathbf{u})$: fit of model bias, the model null hypotheses given a uniform prior (i.e. equal distribution in the regional pool of recruits).</p> <p>ii) $R^2_{KL}(\mathbf{u}, \mathbf{t})$: fit using again a uniform prior but including traits as constraints.</p> <p>iii) $\bar{R}^2_{KL}(\mathbf{m})$: fit using the metacommunity prior but excluding traits as constraints</p> <p>iv) $R^2_{KL}(\mathbf{m}, \mathbf{t})$: fit using the metacommunity prior and including traits as constraints</p> <p>The general form of the R^2_{KL} divergence is calculated by:</p> $5) R^2_{KL} = 1 - \frac{\sum_{k=1}^c \sum_{i=1}^S O_{ik} \ln \left(\frac{O_{ik}}{P_{ik}} \right)}{\sum_{j=k}^c \sum_{i=1}^S O_{ik} \ln \left(\frac{O_{ij}}{Q_{i,0}} \right)}$ <p>With the following parameters:</p> <p>O_{ik} as the observed relative abundances of the i^{th} genus in the k^{th} community,</p> <p>P_{ik} the accompanying predicted values for the specific model of the four solution step as described in the main text and,</p> <p>$Q_{i,0}$ the predicted relative abundances given only the maximum uninformative prior.</p> <p>Further details on the calculation of all separate R^2_{KL} values and accompanying pure trait, pure metacommunity, joint information and biologically unexplained information can be found in the SOM (box S2).</p>
---	---

Box S2. Mathematical description of the Maximum Entropy Formalism for the four-step solution. Left panel shows the necessary ingredients and basic formulation of the Maximum Entropy Formalism. Right side panel shows decomposition of the proportion of total deviance accounted for between observed and predicted relative abundances for each of the four-step solution.

BOX S3

The purpose of using MEF is to decompose the deviance between observed and predicted relative abundances using the four-step solution as described in the main text. The values generated are described below. The R^2_{KL} value is a generalization of the classic R^2 index of maximum likelihood estimation using the Kullback-Leibler index for a non-linear regression including a multinomial error structure [16–18]. In essence, it is a way of measuring the proportion of total deviance accounted for by that specific model from one of the four steps:

$\bar{R}^2_{KL}(\mathbf{u})$: fit of model bias, the model null hypotheses given a uniform prior and permuted traits

$R^2_{KL}(\mathbf{u}, \mathbf{t})$: fit using a uniform prior but including observed traits as constraints

$\bar{R}^2_{KL}(\mathbf{m})$: fit using the metacommunity prior but excluding observed traits as constraints

$R^2_{KL}(\mathbf{m}, \mathbf{t})$: fit using the metacommunity prior and including observed traits as constraints

1) The increase in the explained deviance due to traits can be calculated either by

$$\Delta R^2_{KL}(\mathbf{t}|\boldsymbol{\varphi}) = R^2_{KL}(\mathbf{u}, \mathbf{t}) - \bar{R}^2_{KL}(\mathbf{u})$$

Increase in explained deviance due to traits beyond that due solely to model bias

$$\text{or } \Delta R^2_{KL}(\mathbf{t}|\mathbf{m}) = R^2_{KL}(\mathbf{m}, \mathbf{t}) - \bar{R}^2_{KL}(\mathbf{m})$$

Increase in explained deviance due to traits beyond contributions made by the meta-community

2) The increase in explained deviance due dispersal mass effects via the metacommunity can be calculated by either:

$$\Delta R^2_{KL}(\mathbf{m}|\boldsymbol{\varphi}) = \bar{R}^2_{KL}(\mathbf{m}) - \bar{R}^2_{KL}(\mathbf{u})$$

Increase in explained deviance (if any) due to the metacommunity beyond that due to model bias

$$\text{or } \Delta R^2_{KL}(\mathbf{m}|\mathbf{t}) = R^2_{KL}(\mathbf{m}, \mathbf{t}) - R^2_{KL}(\mathbf{u}, \mathbf{t})$$

Increase in explained deviance due to the meta-community given traits, relative to the explained deviance due only to the traits: i.e. information unique to neutral prior

3) And finally the joint information and the biologically unexplained information:

$$\Delta R^2_{KL}(\mathbf{m}+\mathbf{t}) = \Delta R^2_{KL}(\mathbf{m}|\boldsymbol{\varphi}) - \Delta R^2_{KL}(\mathbf{m}|\mathbf{t}) = \Delta R^2_{KL}(\mathbf{t}|\boldsymbol{\varphi}) - \Delta R^2_{KL}(\mathbf{t}|\mathbf{m})$$

Joint information gain, or increase in explained deviance due to both the metacommunity prior and the constraints based on the traits

$$1 - \Delta R^2_{KL}(\mathbf{m}, \mathbf{t})$$

Biologically unexplained variation

From these values the pure trait, pure metacommunity, joint effect and biologically unexplained variation can be calculated by the following calculations:

$$\text{Pure trait effects: } \Delta R^2_{KL}(\mathbf{t}|\mathbf{m}) / (1 - \bar{R}^2_{KL}(\mathbf{u}))$$

$$\text{Pure metacommunity effects: } \Delta R^2_{KL}(\mathbf{m}|\mathbf{t}) / (1 - \bar{R}^2_{KL}(\mathbf{u}))$$

$$\text{Joint metacommunity and trait effects: } \Delta R^2_{KL}(\mathbf{m}+\mathbf{t}) / (1 - \bar{R}^2_{KL}(\mathbf{u}))$$

$$\text{Unexplained effects: } 1 - \Delta R^2_{KL}(\mathbf{m}, \mathbf{t}) / (1 - \bar{R}^2_{KL}(\mathbf{u}))$$

Box S3. Detailed decomposition of the four-step solution from the MEF. Mathematical description of the decomposition based on the constraints and prior distributions (both uniform and neutral) for each of the steps from the four-step solution to measure the proportion of total deviance accounted for by each specific model from one of the four steps.

FIGURE S1

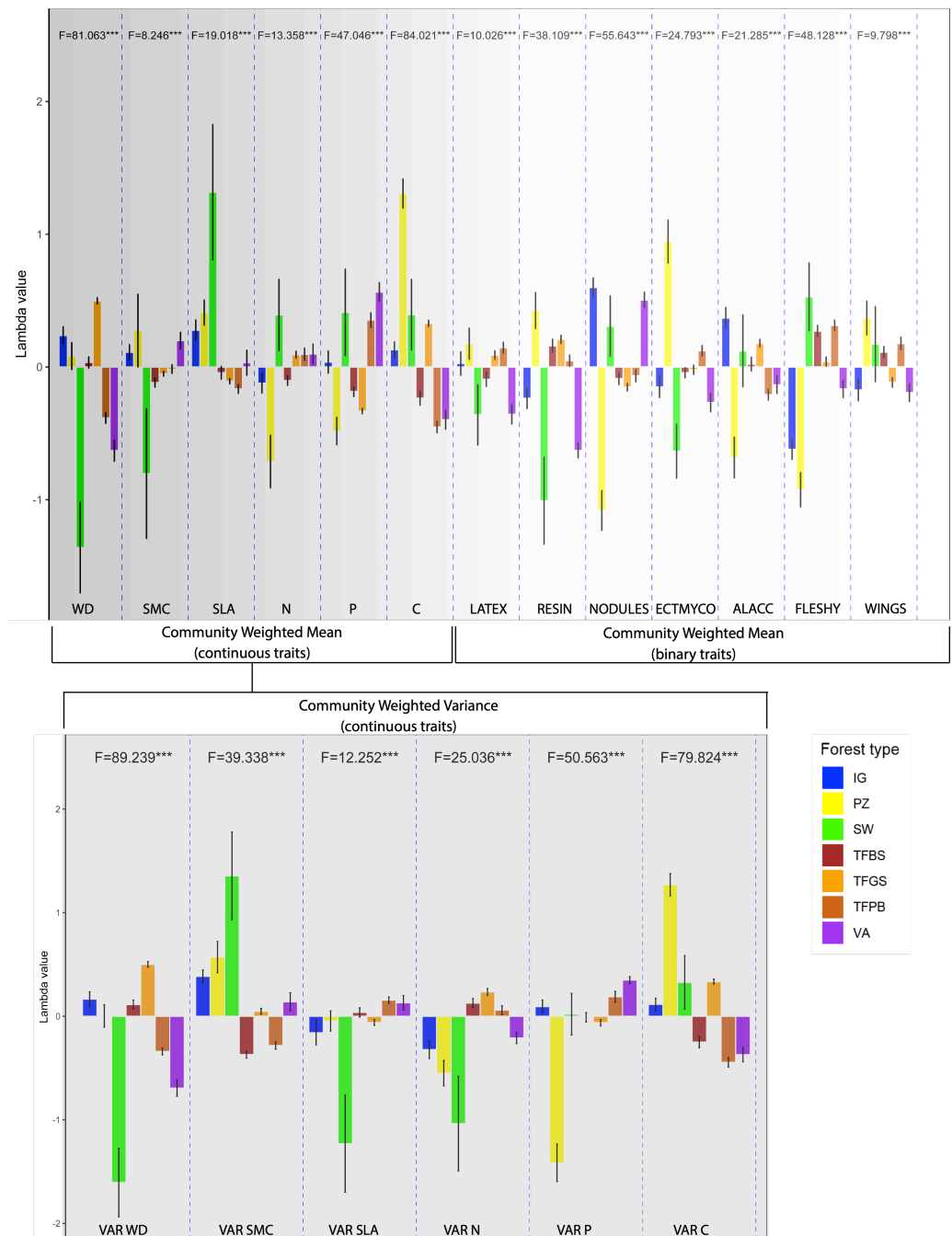


Fig. S1. Mean lambda values with standard error bars for each functional trait and compared between forest types. Forest types are *igapó* (IG), podzol (PZ), swamp (SW), Brazilian shield *terra firme* (TFBS), Guiana Shield *terra firme* (TFGS), Pebas *terra firme* (TFPB) and *várzea* (VA). Differences were tested with a one way analysis of variance with significance levels corresponding to: ns non-significant, * $p < .05$, ** $p < .01$ and *** $p < .001$. Traits used were wood density (WD), seed mass class (SMC), specific leaf area (SLA), nitrogen (N), phosphorus (P) and carbon (C) leaf content with the prefix of VAR for the variance of continuous traits. Latex, Resin, Nodules, Ectomycorrhiza (EctoMyco), the ability to accumulate aluminum (AlAcc), and the presence/absence of fleshy fruits (Fleshy) and winged seeds (Wings) were all binary traits.

FIGURE S2

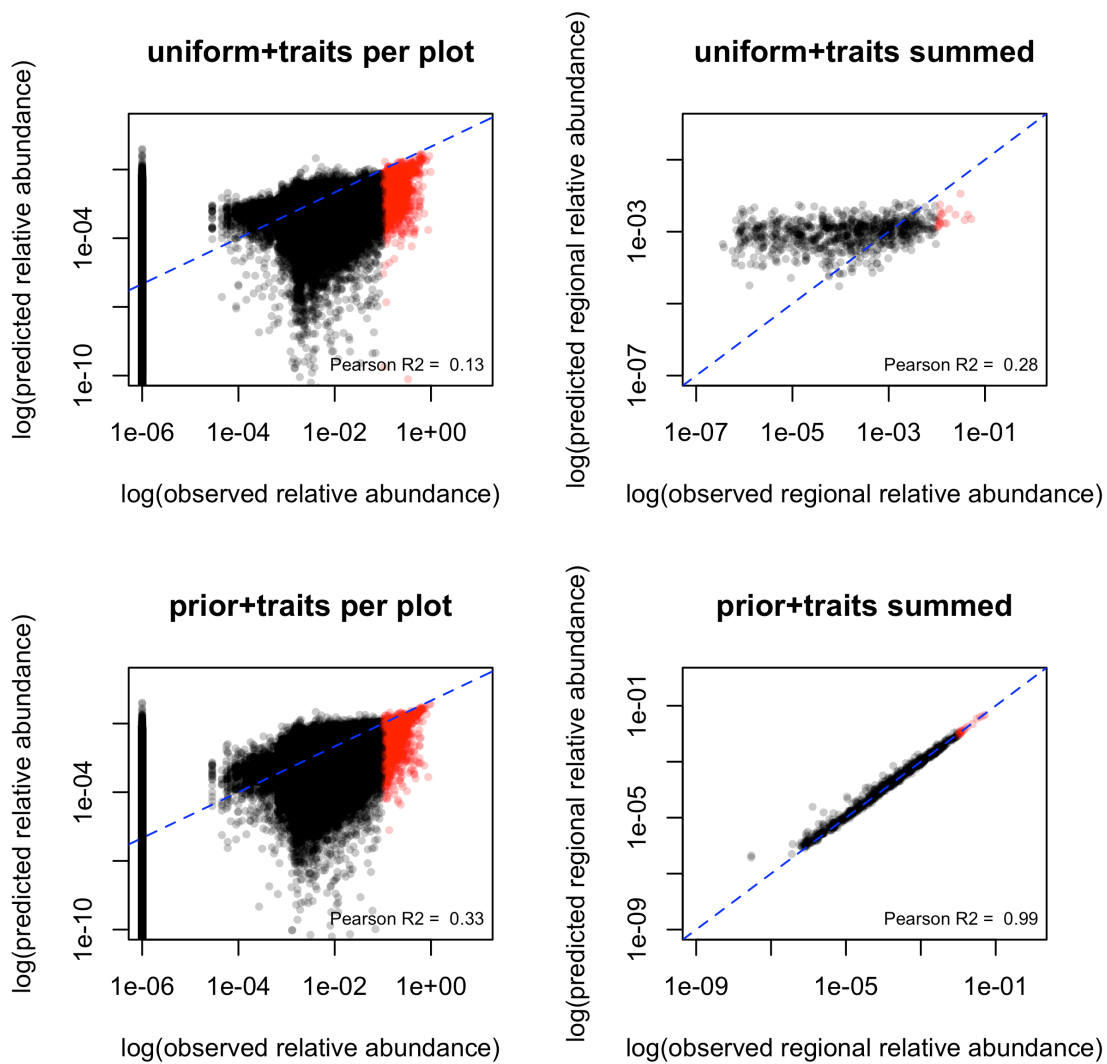


Fig. S2. Observed relative abundances for each genus in all plots plotted against predicted relative abundance per plot (left) and summed (right) using only the traits as constraints in combination with a uniform prior (top) or the hybrid model using both traits and the metacommunity relative abundance as prior (bottom) on a log-log scale. Top figures show predictions using only a uniform prior, left separate for all plots and right for all genera summed over all plots. Bottom figures show predictions using the regional prior, again separate for all plots and genera (left) and summed over all plots for each genus (right). Red points indicate taxa with observed relative abundances over 1e-1. Lines show the $x=y$ prediction and R^2 values correspond to the Pearson's correlation coefficient. Reported R^2 value is equal to the Pearson correlation coefficient (R) between the observed and predicted relative abundances defined as one minus the ratio of the error sum of squares to the total sum of squares.

FIGURE S3

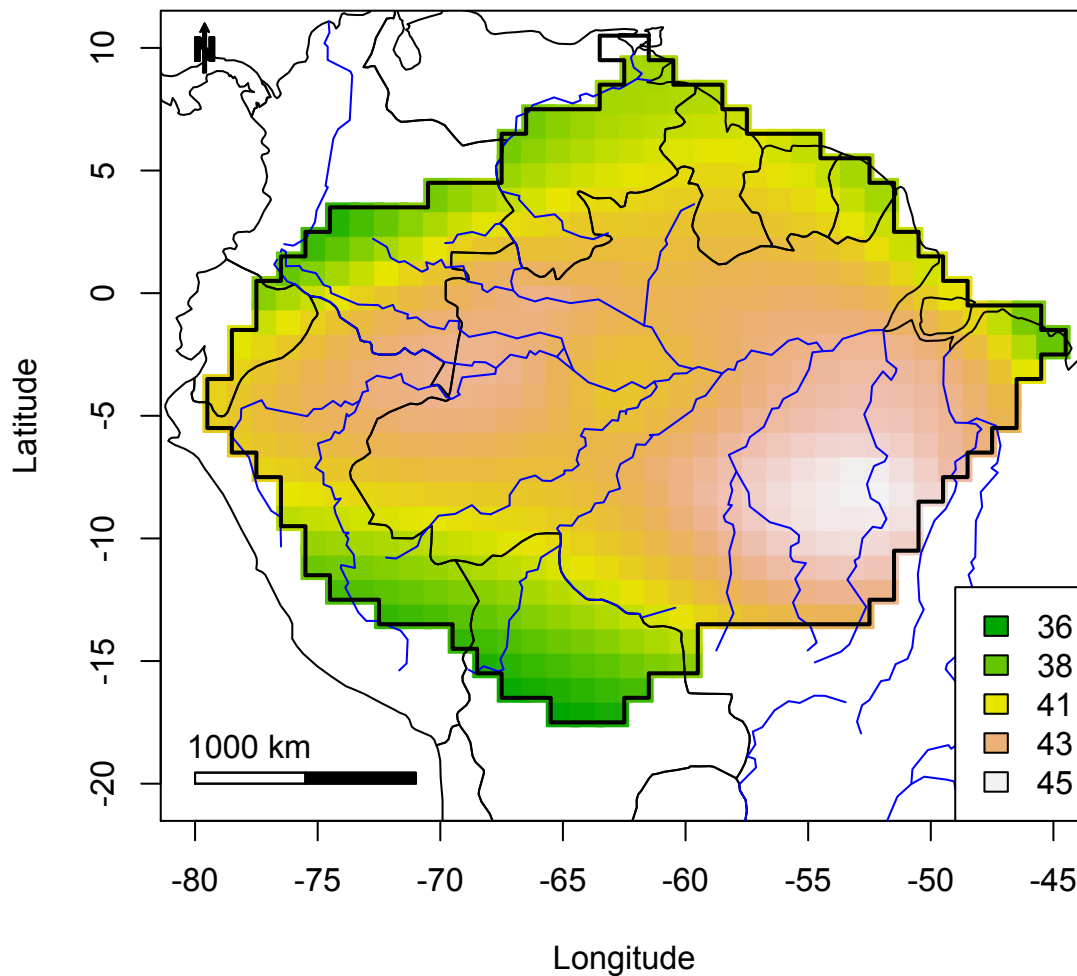


Fig. S3. Spatial gradient in pure trait relative to pure metacommunity effect. Map showing the ratio between the pure metacommunity effect and the pure trait effects for each plot. Ratio was calculated per plot by dividing the pure metacommunity effect by $(10 \times \text{pure trait effects} + 1)$. Values for projection on the map using a loess regression were multiplied by 1000 to allow clearer differentiation. Squares show the predictions from loess regression (color depending on value). Map shows interior of the Amazon having weaker trait effects relative to metacommunity effects whereas on the edges of the Amazon this pattern is reversed.

FIGURE S4

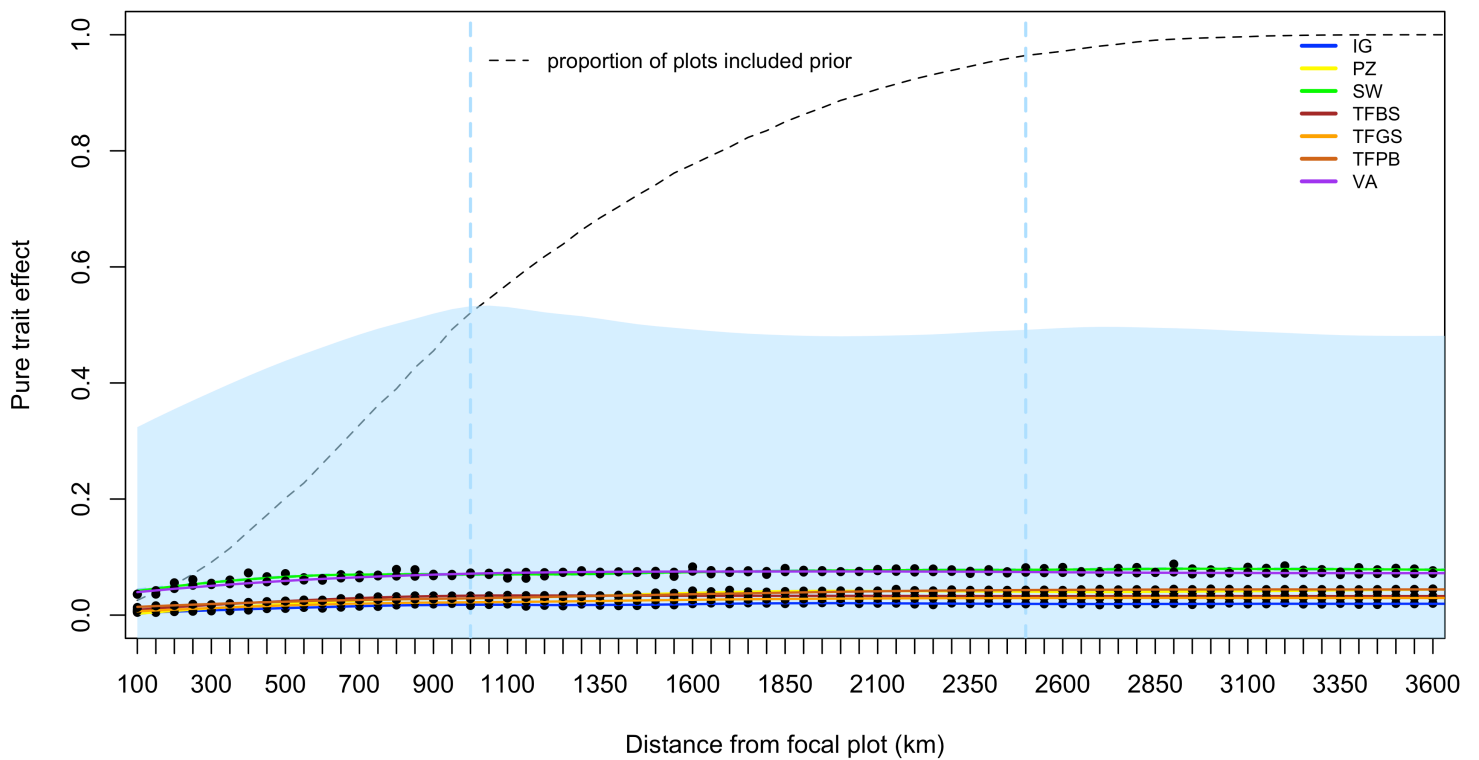


Fig. S4. Distance decay of pure trait effect for each forest type separately and the overall dataset. X-axis represents the radius of the metacommunity prior; i.e. the first 100 km consists of just a few plots and at 3800 km all plots are taken into account. Y-axis represent the pure trait effect, i.e. the increase in explained deviance due to traits beyond contributions made by the meta-community and relative to the model bias (see also Box S2). Colors indicate the different forest types with abbreviations as in main text. Lines indicate the predictions following from the loess regression based on all points. Blue vertical lines indicate the 1000 and 2500 km boundary points. Blue shading reflects maximum values for that distance of the whole dataset.

FIGURE S5

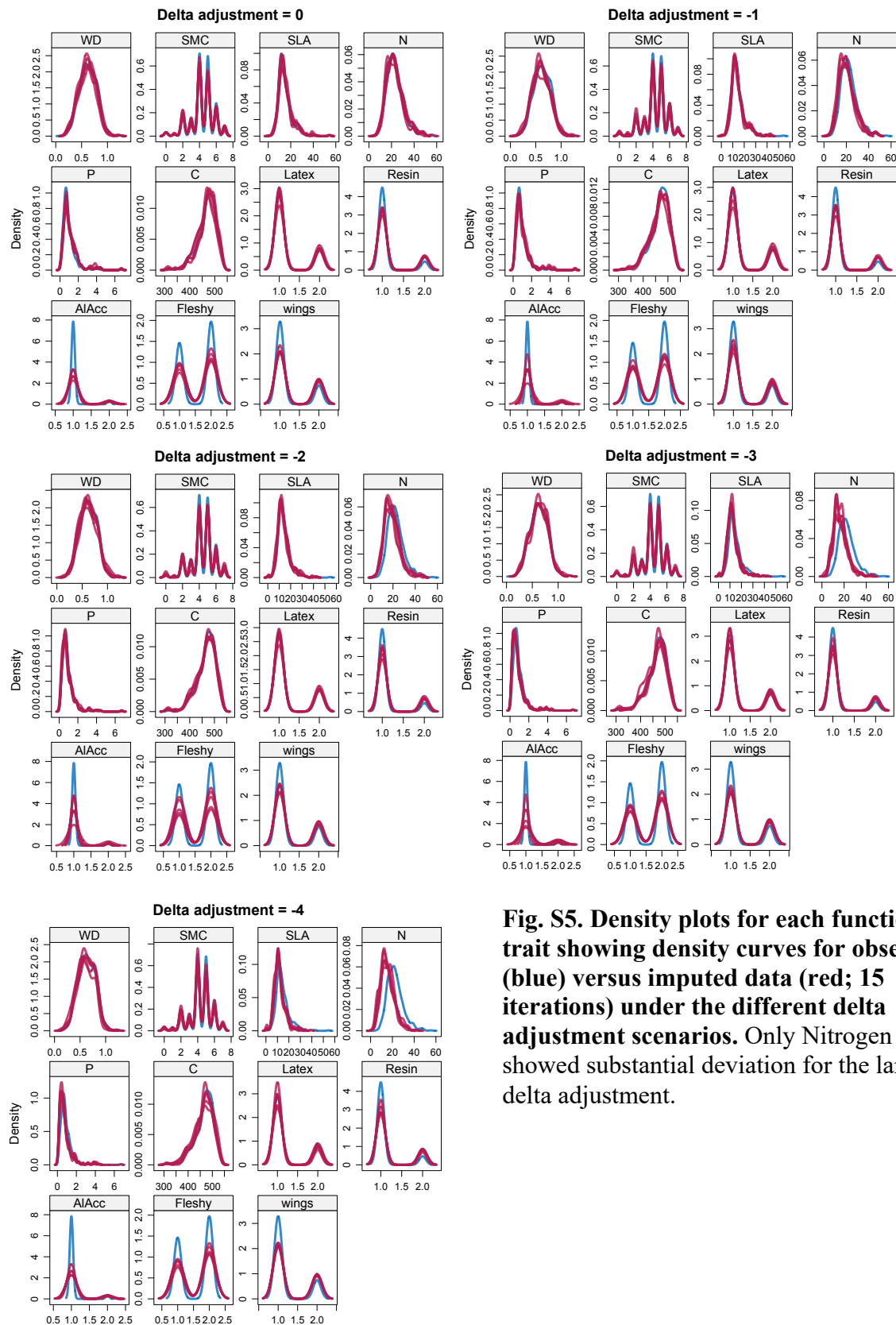


Fig. S5. Density plots for each functional trait showing density curves for observed (blue) versus imputed data (red; 15 iterations) under the different delta adjustment scenarios. Only Nitrogen showed substantial deviation for the larger delta adjustment.

FIGURE S6

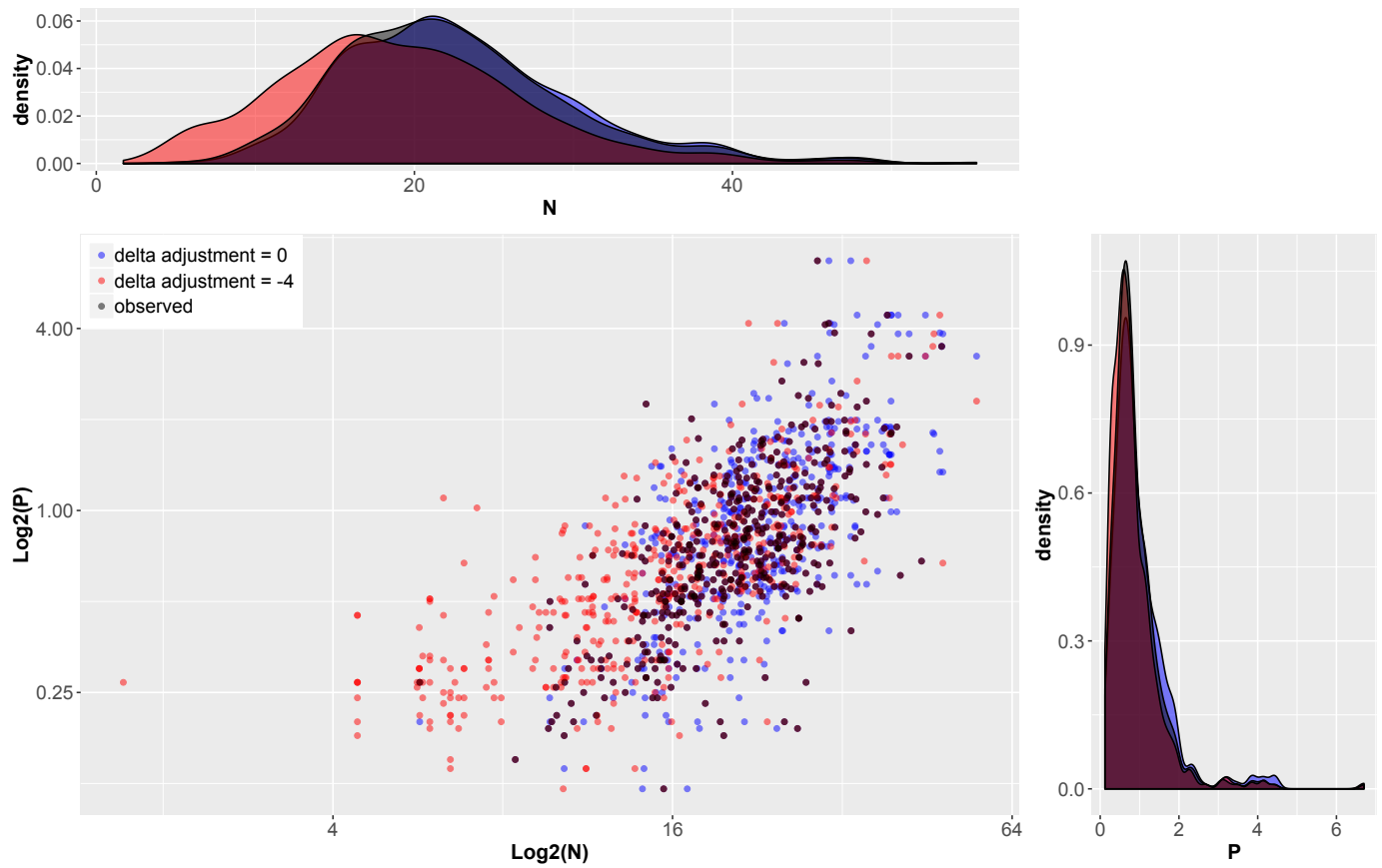


Fig. S6. Scatterplot of observed vs. imputed leaf nitrogen and phosphorus content under different delta adjustment scenarios. As Leaf Nitrogen content showed substantial deviation in the larger delta adjustment scenario it was here plotted versus Leaf Phosphorus content, which showed hardly any deviation.

FIGURE S7

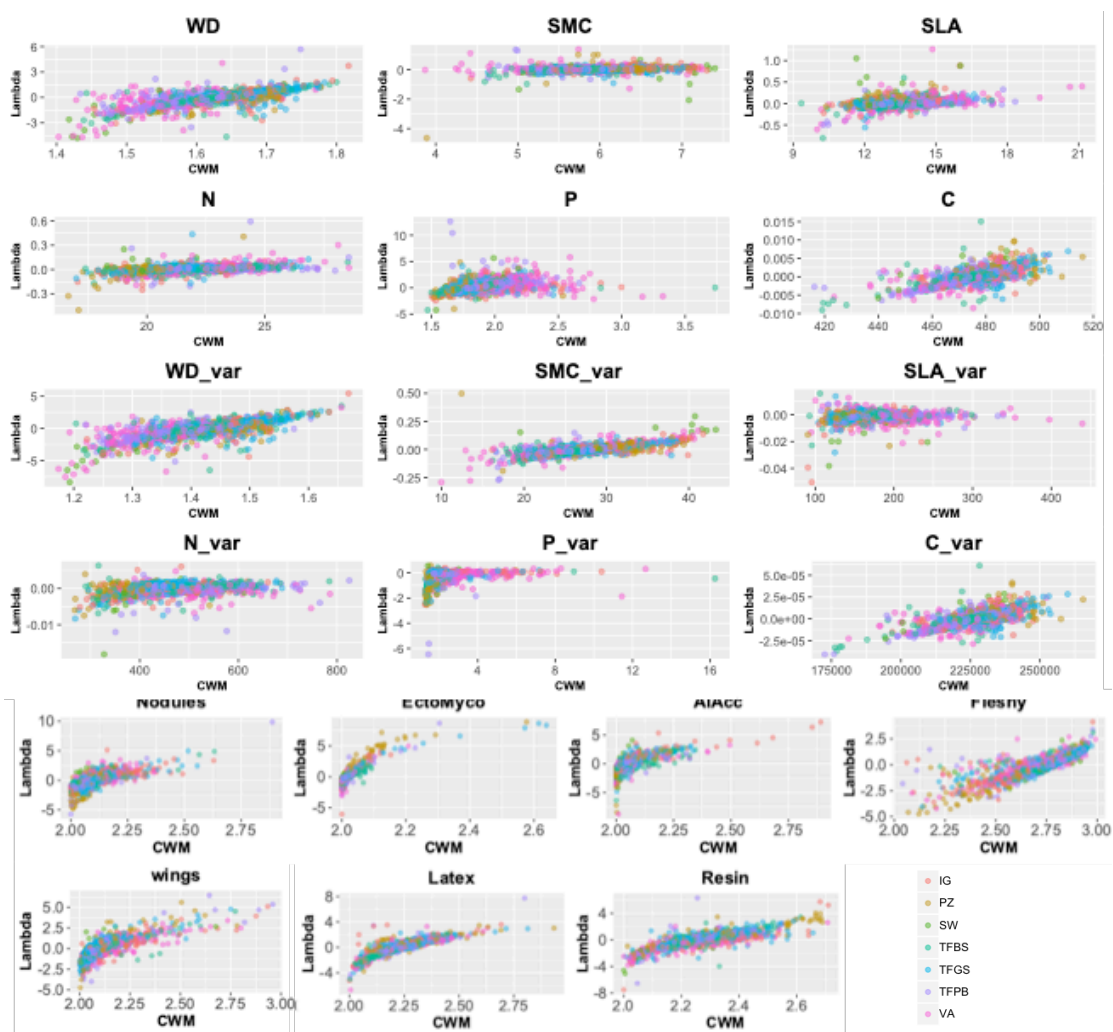


Fig. S7 Scatterplots of the CWM values against lambda values, colored according to forest type. Titles are abbreviations for functional traits as used in the main text. Plots show in some cases these are clearly correlated (e.g. wood density, seedmassclass and C) but for many others not (e.g. SLA, Latex or Nodules).

FIGURE S8

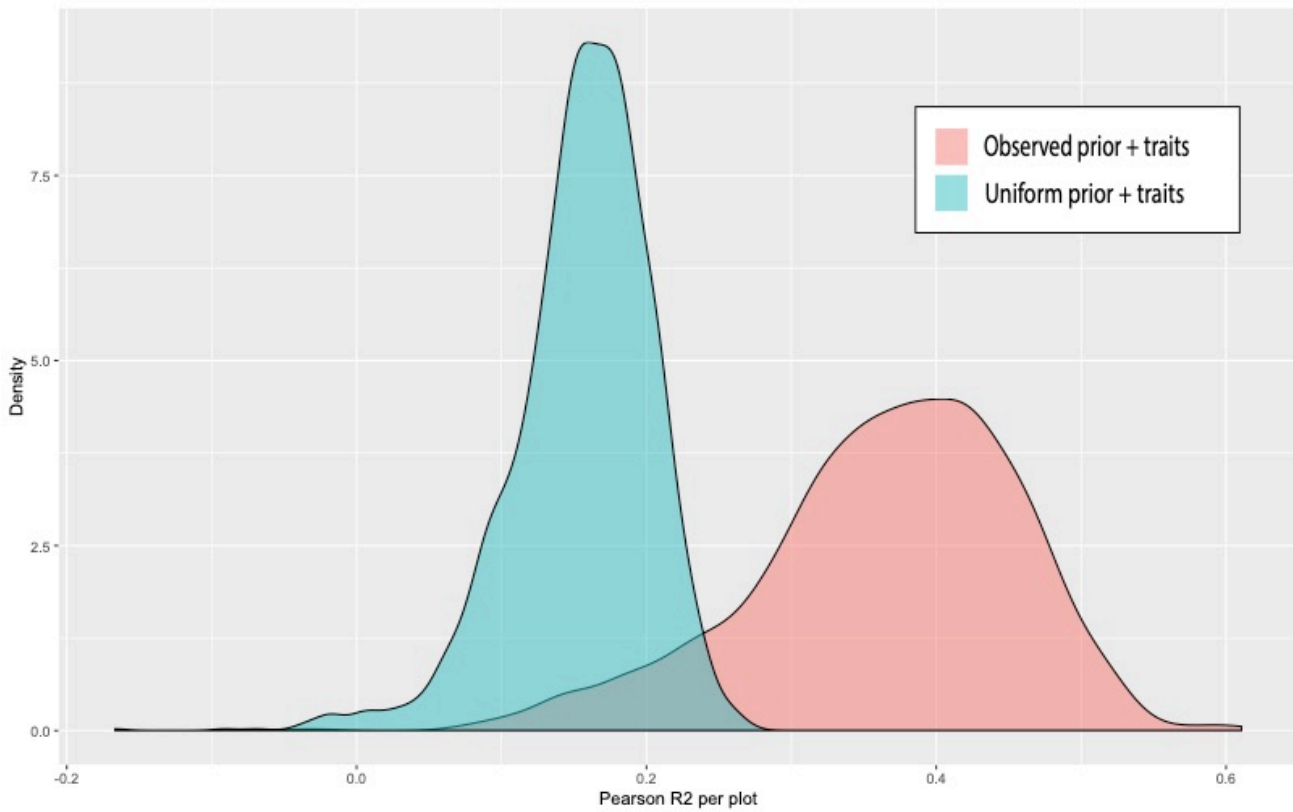


Fig. S8 Density plot of the per-plot-Pearson correlation coefficient between predicted relative abundances of each genus. Models either used a uniform prior and functional traits (blue) or the actual observed prior and functional traits (red), results show a large increase in accuracy for the latter.

FIGURE S9

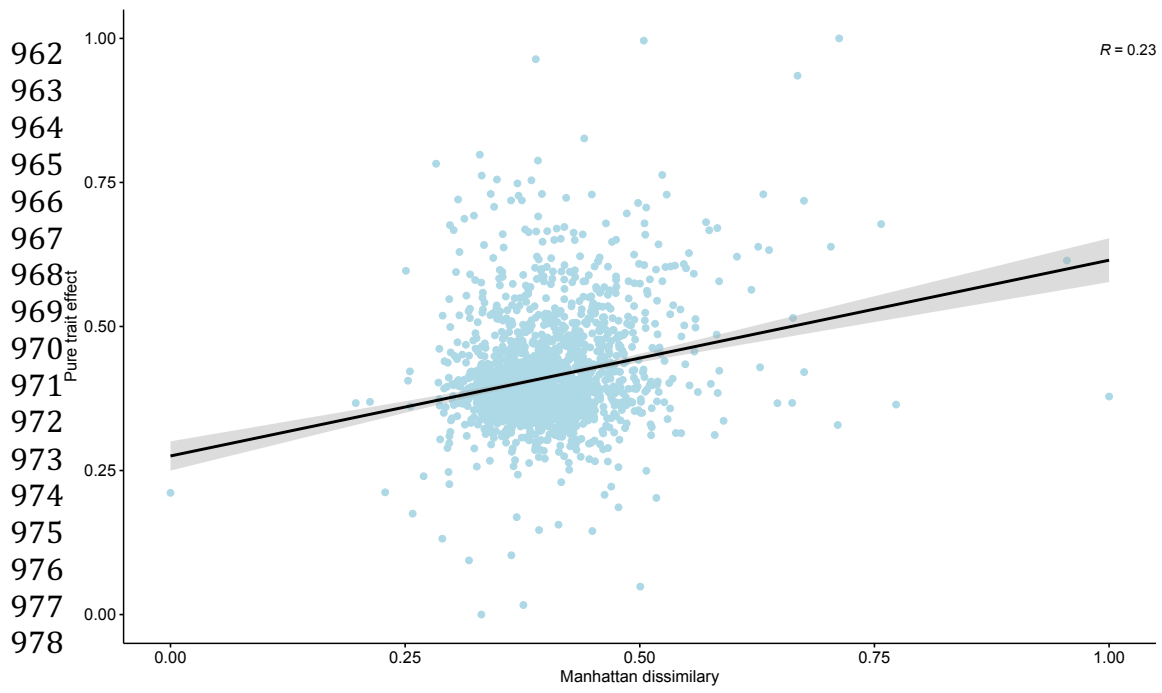


FIGURE S10

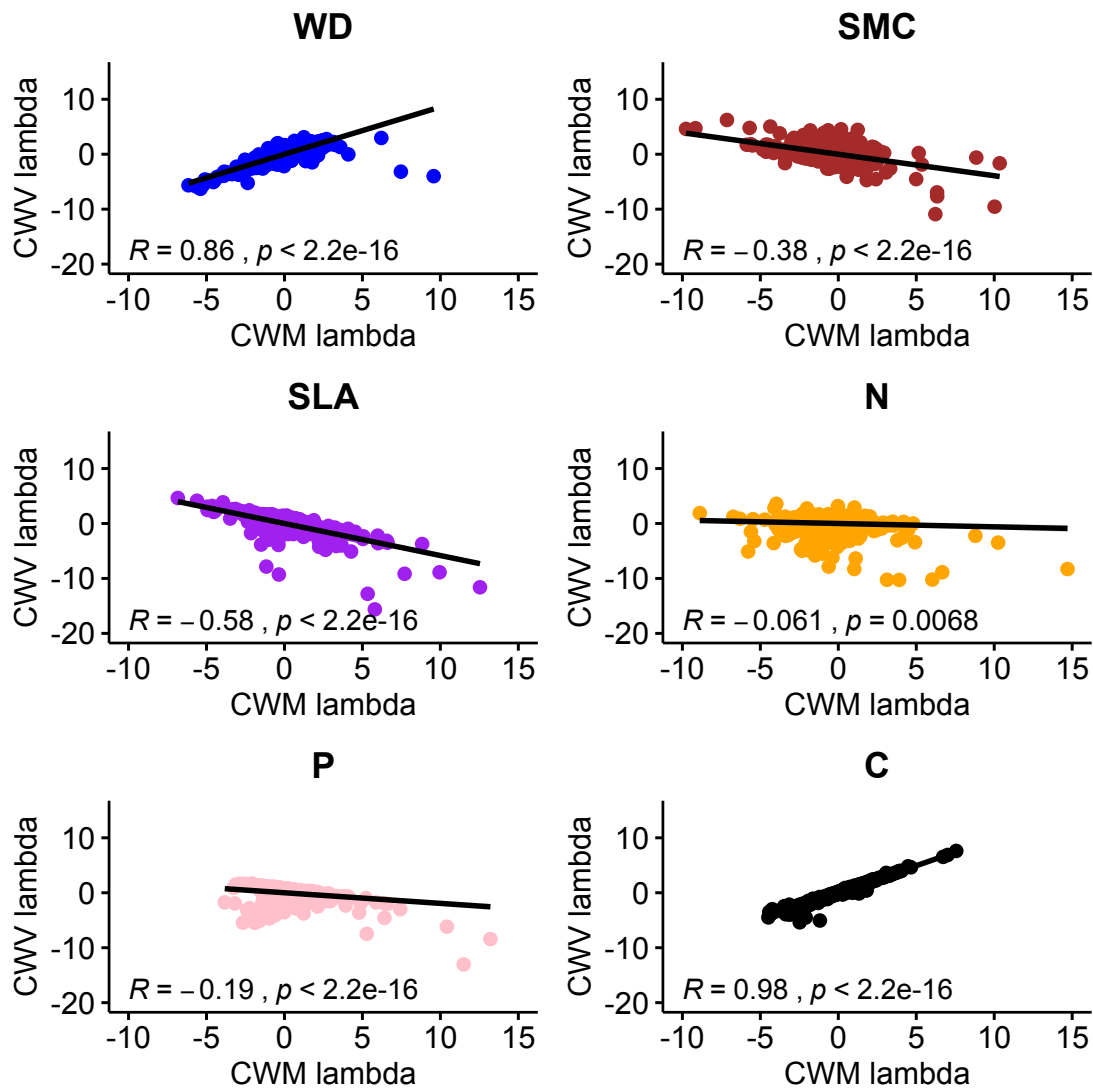


TABLE S1

Table S1. Decomposition of results from the various maximum entropy models, combined and separated by forest type (PZ podzol, IG igapó, VA várzea, SW Swamp, TF terra firme with subregions BS (Brazilian Shield), GS (Guyana Shield) and PB (Pebas formation)). Top rows indicate estimated proportions (R^2_{KL}) of total information reflective of variation in local relative abundance explained for by the various models. Middle rows indicate the specific information gain from any one of the used models relative to the model bias. Bottom rows show the actual effects of traits, the metacommunity and the joint information relative to the model bias.

Explained proportions	Forest types							Combined
	PZ	VA	IG	SW	TFB S	TFGS	TFPB	
$\bar{R}^2_{KL}(u)$ <i>model bias fit</i>	0.20	0.14	0.15	0.23	0.11	0.11	0.09	0.12
$\bar{R}^2_{KL}(m)$ <i>pure neutral model fit</i>	0.54	0.53	0.54	0.57	0.56	0.63	0.56	0.58
$R^2_{KL}(u,t)$ <i>pure trait model fit</i>	0.36	0.25	0.23	0.34	0.20	0.24	0.21	0.23
$R^2_{KL}(m,t)$ <i>hybrid model fit</i>	0.60	0.59	0.56	0.60	0.60	0.66	0.60	0.62
Increase in explained deviance								
$\Delta R^2_{KL}(m \phi)$ <i>metacommunity effect beyond model bias</i>	0.34	0.39	0.39	0.34	0.45	0.52	0.47	0.45
$\Delta R^2_{KL}(t \phi)$ <i>trait effect beyond model bias</i>	0.15	0.11	0.08	0.11	0.10	0.13	0.11	0.11
$\Delta R^2_{KL}(t m)$ <i>trait effect beyond metacommunity effect</i>	0.06	0.06	0.02	0.03	0.04	0.03	0.05	0.04
$\Delta R^2_{KL}(m t)$ <i>metacommunity effect relative to trait effects</i>	0.24	0.34	0.33	0.27	0.40	0.43	0.40	0.38
$\Delta R^2_{KL}(m+t)$ <i>joint effect of metacommunity and traits</i>	0.09	0.05	0.06	0.07	0.06	0.10	0.07	0.07
$1 - \Delta R^2_{KL}(m,t)$ <i>unexplained effects</i>	0.40	0.41	0.44	0.40	0.40	0.34	0.40	0.38
Biologically relevant information								
Pure trait effect <i>Information from traits, relative to bias</i>	0.08	0.07	0.03	0.05	0.05	0.04	0.05	0.05
Pure metacommunity effect <i>Information from metacommunity, relative to bias</i>	0.30	0.39	0.38	0.33	0.45	0.48	0.44	0.43
Joint effect <i>Information from joint effect, relative to bias</i>	0.12	0.06	0.07	0.10	0.06	0.11	0.07	0.08
Unexplained information <i>Left over information not explained, relative to bias</i>	0.50	0.48	0.52	0.52	0.45	0.38	0.44	0.44

TABLE S2

Table S2. Summary statistics overview for the linear models of the various scenarios under the delta adjustment technique as described in the main text.

Rows indicate the different delta adjustments used with the columns representing the standard summary statistics of the linear model comparing the imputed versus observed trait values. Results showed similar patterns with each imputation scenario, indicating a robust imputation procedure.

Scenario	Summary statistics linear model $lm(\text{imputed} \sim \text{observed})$						
	Intercept	Std. error	T value	Pr.	R ²	Adj. R ²	Sigma
<i>Delta 0</i>	-.33	.07	-4.85	1.42e-06	.32	.33	.63
<i>Delta -2.5</i>	-.34	.06	-5.93	4.31e-09	.37	.37	.58
<i>Delta -5</i>	-.16	.04	-3.57	3.77e-04	.40	.40	.53
<i>Delta -7.5</i>	.64	.09	-7.33	5.18e-13	.42	.42	.88
<i>Delta -10</i>	0.09	.04	3.17	1.57e-03	.47	.47	.48

979

980

981

982 S-A Ecological interpretation of the MEF results

983

984 A number of functional traits associated with low nutrient conditions (e.g.
985 ectomycorrhiza) and life history strategies suited for protection against herbivores
986 (e.g. latex and high leaf C content) were clearly positively associated with abundance
987 in nutrient poor environments (podzols) in terms of community weighted mean
988 values, indicated by the positive lambda values. In contrast, community weighted
989 means for fleshy fruits and high leaf N and P content were negatively associated with
990 abundance on these soils. Nodulation was also negatively associated with abundance
991 on poor soils, supporting earlier results [1]. The ability to accumulate aluminium was
992 positively associated with abundance on soils commonly associated with higher
993 aluminium content such as *igapó* (strong positive effects) and *terra firme* soils (a
994 minor, yet positive effect). In contrast, it was strongly negatively associated with
995 abundance for podzol, *várzea* and swamp forests. Traits such as SLA or winged fruits
996 also showed strong patterns dependent on forest type.

997

998 Signals of quantitative environmental selection were found to be highest for podzol
999 forests, whereas its counterpart in the form of the dispersal mass effect from the
1000 regional pool of genera had the lowest value. Podzol forests, having extremely
1001 nutrient poor soils could reflect a much stronger selective environment than any of the
1002 other forest types. *Terra firme* forests, presumably reflective of a less strong selective
1003 environment in terms of resource availability, showed the opposite, with less than half
1004 of the pure trait effect in comparison with podzol forests (even when rarefied to
1005 accommodate for different sample sizes). Traits associated with protection against
1006 herbivores such as latex [2] and high leaf carbon content showed higher values
1007 associated with greater abundance and overall lower variance on podzol soils,
1008 whereas traits indicative of investment in growth and photosynthetic ability such as
1009 high foliar concentrations of P and N [3] showed strong negative associations on
1010 nutrient poor soils for both community weighted means and variance. The ability to
1011 accumulate aluminium was also strongly positively associated with relative
1012 abundance on the more nutrient but also often aluminium enriched soils of *terra firme*
1013 and in some cases aluminium rich *igapó* forests. Lambda values also showed strong
1014 negative lambda values for wood density in swamp and forests in both community
1015 weighted means and variance, fitting high tree mortality and many individuals
1016 belonging to pioneer species in especially the western Amazonian swamp forests.
1017 *Várzea* and Pebas *terra firme* forests showed a similar response. As the Pebas consists
1018 mainly of Andean sediments it has higher nutrient content, promoting lower wood
1019 density, supported by our results whereas *várzea* forests are also often flooded. There
1020 were also traits that showed no specific (strong) signal of selection on certain forest
1021 types (either positive or negative), such as latex on *igapó* and ectomycorrhiza on
1022 *várzea* (see Fig. 1 for all lambda values). Plotting lambda values for CWM and CWV
1023 constraints of the continuous traits showed WD and C were both strongly positively
1024 correlated indicating strong directional selection for lower trait values accompanied
1025 by a reduction in trait variance. SMC, SLA and leaf P content, however, showed a
1026 negative correlation with higher lambda values for CWM values associated with
1027 lower trait variance (Figure S11). None of the traits showed a reduction in variance
1028 without a change in the CWM, suggesting directional selection is more likely than
1029 stabilizing selection, even though the overall information yield remains low.
1030 Interestingly, terra firme forests in general showed the smallest lambda values overall
1031 (positive or negative).

1032

1033 This may be indicative of either more pronounced demographic stochasticity or
1034 ecological drift eliminating the association between traits and relative abundance.
1035 Lower effects of selection in general or more (random) variation due to the larger
1036 species pool in comparison with other forest types, however, could also be the result
1037 of mixing heterogeneous microenvironments into a single environmental class.
1038 Support for such heterogeneity within terra firme forests having influence on
1039 distribution of functional traits on valleys or plateaus has recently been found [4]. In
1040 addition, natural but also anthropogenic [5] disturbance history affects biotic
1041 community composition and can lead to changes in tree community through time,
1042 blurring relationships between traits and relative abundances. It should further be
1043 noted that, although for terra firme forests we were able to make a distinction by
1044 subregion, true within forest type heterogeneity was not taken into account. This
1045 might cause an underestimation of the deterministic effect but as of yet cannot be
1046 corrected for on this scale and is worth to be investigated in future studies. In
1047 addition, podzol forests have a smaller connected surface area and accompanying
1048 smaller number of genera in comparison with terra firme forests, adding to the
1049 calculated stronger trait effects [6,7]. When more detailed understanding and
1050 knowledge of these functional traits would be provided, this would most likely
1051 increase the explanatory power of the MEF. The fact, however, that we do not have a
1052 very specific knowledge of these interactions and specific traits is precisely the reason
1053 why the MEF can provide additional insight.

1054
1055 It should be noted that for species level analyses any micro environmental gradients
1056 might prove to also show (stronger) selection at local scales [8,9], as it has been
1057 shown that most variation in community composition, due to selection in regard to
1058 habitat filtering and niche conservatism, is found at lower taxonomic levels, such as
1059 between species within genera [10,11]. In contrast, theoretically it has been shown
1060 and tested that immigration numbers are actually very robust across taxonomic
1061 scales [12], validating our results of the metacommunity importance using genus
1062 level taxonomy. Spatial patterns of metacommunity effects, showing shallowest
1063 declines in the centre, also support the suggestion that high diversity of the
1064 Amazonian interior could be explained by influx of recruits due to large (overlapping)
1065 ranges. This mid-domain effect [13], however, would also predict lower species
1066 richness for the edges due to lower range overlap, assuming a closed community. This
1067 is not the case, as there is a strong species richness gradient from West (rich) to
1068 Eastern Amazonian forests (poor) [14]. The lower metacommunity effect for the
1069 edges then is most likely not due to less absolute influx of genera, but rather less
1070 influx from the Amazonian tree community. Influx from the species-rich Andes could
1071 account for the high diversity [15], yet low Amazonian metacommunity effect for
1072 Western Amazonian forests. In contrast, South Eastern parts of Amazonia receive
1073 influx from tree species-poor biomes (i.e. the Cerrado) resulting in lower diversity but
1074 also low metacommunity effect for Amazonian trees in this region.
1075

1076 **REFERENCES SUPPLEMENTARY MATERIAL**

1077

- 1078 [1] H. ter Steege, N. C. a Pitman, O. L. Phillips, J. Chave, D. Sabatier, A. Duque,
 1079 J.-F. Molino, M.-F. Prévost, R. Spichiger, H. Castellanos, P. von Hildebrand,
 1080 and R. Vásquez, *Continental-Scale Patterns of Canopy Tree Composition and*
 1081 *Function across Amazonia.*, *Nature* **443**, 444 (2006).
- 1082 [2] A. A. Agrawal, *Macroevolution of Plant Defense Strategies*, *Trends Ecol. Evol.*
 1083 **22**, 103 (2007).
- 1084 [3] J. C. Ordoñez, P. M. Van Bodegom, J. P. M. Witte, I. J. Wright, P. B. Reich,
 1085 and R. Aerts, *A Global Study of Relationships between Leaf Traits, Climate*
 1086 *and Soil Measures of Nutrient Fertility*, *Glob. Ecol. Biogeogr.* **18**, 137 (2009).
- 1087 [4] R. S. Oliveira, F. R. C. Costa, E. van Baalen, A. de Jonge, P. R. Bittencourt, Y.
 1088 Almanza, F. de V. Barros, E. C. Cordoba, M. V. Fagundes, S. Garcia, Z. T. T.
 1089 M. Guimaraes, M. Hertel, J. Schietti, J. Rodrigues-Souza, and L. Poorter,
 1090 *Embolism Resistance Drives the Distribution of Amazonian Rainforest Tree*
 1091 *Species along Hydro-Topographic Gradients*, *New Phytol.* (2018).
- 1092 [5] C. Levis, B. M. Flores, P. A. Moreira, B. G. Luize, R. P. Alves, J. Franco-
 1093 Moraes, J. Lins, E. Konings, M. Peña-Claros, F. Bongers, F. R. C. Costa, and
 1094 C. R. Clement, *How People Domesticated Amazonian Forests*, *Front. Ecol.*
 1095 *Evol.* **5**, (2018).
- 1096 [6] J. E. Guevara, G. Damasco, C. Baraloto, P. V. A. Fine, M. C. Peñuela, C.
 1097 Castilho, A. Vincentini, D. Cárdenas, F. Wittmann, N. Targhetta, O. Phillips, J.
 1098 Stropp, I. Amaral, P. Maas, A. Monteagudo, E. M. Jimenez, R. Thomas, R.
 1099 Brienen, Á. Duque, W. Magnusson, C. Ferreira, E. Honorio, F. de Almeida
 1100 Matos, F. R. Arevalo, J. Engel, P. Petronelli, R. Vasquez, and H. ter Steege,
 1101 *Low Phylogenetic Beta Diversity and Geographic Neo-Endemism in*
 1102 *Amazonian White-Sand Forests*, *Biotropica* **48**, 34 (2016).
- 1103 [7] H. ter Steege, N. C. A. Pitman, D. Sabatier, C. Baraloto, R. P. Salomão, J. E.
 1104 Guevara, O. L. Phillips, C. V Castilho, W. E. Magnusson, J.-F. Molino, A.
 1105 Monteagudo, P. Núñez Vargas, J. C. Montero, T. R. Feldpausch, E. N. H.
 1106 Coronado, T. J. Killeen, B. Mostacedo, R. Vasquez, R. L. Assis, J. Terborgh, F.
 1107 Wittmann, A. Andrade, W. F. Laurance, S. G. W. Laurance, B. S. Marimon,
 1108 B.-H. Marimon, I. C. Guimarães Vieira, I. L. Amaral, R. Brienen, H.
 1109 Castellanos, D. Cárdenas López, J. F. Duivenvoorden, H. F. Mogollón, F. D. de
 1110 A. Matos, N. Dávila, R. García-Villacorta, P. R. Stevenson Diaz, F. Costa, T.
 1111 Emilio, C. Levis, J. Schietti, P. Souza, A. Alonso, F. Dallmeier, A. J. D.
 1112 Montoya, M. T. Fernandez Piedade, A. Araujo-Murakami, L. Arroyo, R.
 1113 Gribel, P. V. A. Fine, C. A. Peres, M. Toledo, G. A. Aymard C, T. R. Baker, C.
 1114 Cerón, J. Engel, T. W. Henkel, P. Maas, P. Petronelli, J. Stropp, C. E. Zartman,
 1115 D. Daly, D. Neill, M. Silveira, M. R. Paredes, J. Chave, D. de A. Lima Filho, P.
 1116 M. Jørgensen, A. Fuentes, J. Schöngart, F. Cornejo Valverde, A. Di Fiore, E.
 1117 M. Jimenez, M. C. Peñuela Mora, J. F. Phillips, G. Rivas, T. R. van Andel, P.
 1118 von Hildebrand, B. Hoffman, E. L. Zent, Y. Malhi, A. Prieto, A. Rudas, A. R.
 1119 Ruschell, N. Silva, V. Vos, S. Zent, A. A. Oliveira, A. C. Schutz, T. Gonzales,
 1120 M. Trindade Nascimento, H. Ramirez-Angulo, R. Sierra, M. Tirado, M. N.
 1121 Umaña Medina, G. van der Heijden, C. I. A. Vela, E. Vilanova Torre, C.
 1122 Vriesendorp, O. Wang, K. R. Young, C. Baider, H. Balslev, C. Ferreira, I.
 1123 Mesones, A. Torres-Lezama, L. E. Urrego Giraldo, R. Zagt, M. N. Alexiades,
 1124 L. Hernandez, I. Huamantupa-Chuquimaco, W. Milliken, W. Palacios Cuenca,
 1125 D. Pauletto, E. Valderrama Sandoval, L. Valenzuela Gamarra, K. G. Dexter, K.

- 1126 Feeley, G. Lopez-Gonzalez, M. R. Silman, S. P. Hubbell, F. He, R. Condit, L.
1127 Borda-de-Agua, J. Kellner, H. Ter Steege, G. A. Black, T. H. Dobzhansky, C.
1128 Pavan, J. M. Pires, T. Dobzhansky, G. A. Black, M. J. G. Hopkins, M. J.
1129 Costello, R. M. May, N. E. Stork, P. Haripersaud, H. ter Steege, J.-J. de
1130 Granville, H. Chevillotte, M. Hoff, D. P. Bebber, M. A. Carine, J. R. Wood, A.
1131 H. Wortley, D. J. Harris, G. T. Prance, G. Davidse, J. Paige, T. D. Pennington,
1132 N. K. Robson, R. W. Scotland, B. J. McGill, R. S. Etienne, J. S. Gray, D.
1133 Alonso, M. J. Anderson, H. K. Benecha, M. Dornelas, B. J. Enquist, J. L.
1134 Green, F. He, A. H. Hurlbert, A. E. Magurran, P. A. Marquet, B. A. Maurer, A.
1135 Ostling, C. U. Soykan, K. I. Ugland, E. P. White, R. J. Warren, D. K. Skelly,
1136 O. J. Schmitz, M. A. Bradford, N. C. A. Pitman, J. W. Terborgh, M. R. Silman,
1137 P. N. V, D. A. Neill, C. E. Cerón, W. A. Palacios, M. Aulestia, N. C. A.
1138 Pitman, M. R. Silman, J. W. Terborgh, F. D. Lozano, M. W. Schwartz, M. W.
1139 Schwartz, D. Simberloff, J. E. Richardson, R. T. Pennington, T. D. Pennington,
1140 P. M. Hollingsworth, T. L. Couvreur, F. Forest, W. J. Baker, S. Cavers, C. W.
1141 Dick, D. H. Janzen, S. A. Mangan, S. A. Schnitzer, E. A. Herre, K. M. Mack,
1142 M. C. Valencia, E. I. Sanchez, J. D. Bever, W. Balée, D. G. Campbell, C.
1143 Levis, P. F. de Souza, J. Schietti, T. Emilio, J. L. P. V. Pinto, C. R. Clement, F.
1144 R. C. Costa, D. A. Posey, E. Montoya, V. Rull, N. D. Stansell, M. B. Abbott, S.
1145 Nogué, B. W. Bird, W. A. Díaz, C. Gomez-Navarro, C. Jaramillo, F. Herrera,
1146 S. L. Wing, R. Callejas, C. H. McMichael, D. R. Piperno, M. B. Bush, M. R.
1147 Silman, A. R. Zimmerman, M. F. Raczka, L. C. Lobato, H. ter Steege, P. P.
1148 Haripersaud, O. S. Bánki, F. Schieving, S. J. Phillips, R. P. Anderson, R. E.
1149 Schapire, S. J. Phillips, M. Dudik, C. A. Quesada, J. Lloyd, L. O. Anderson, N.
1150 M. Fyllas, M. Schwarz, C. I. Czimczik, B. Rollet, H. ter Steege, N. C. Pitman,
1151 O. L. Phillips, J. Chave, D. Sabatier, A. Duque, J. F. Molino, M. F. Prévost, R.
1152 Spichiger, H. Castellanos, P. von Hildebrand, R. Vásquez, P. M. Fearnside, D.
1153 Mouillot, D. R. Bellwood, C. Baraloto, J. Chave, R. Galzin, M. Harmelin-
1154 Vivien, M. Kulbicki, S. Lavergne, S. Lavorel, N. Mouquet, C. E. Paine, J.
1155 Renaud, W. Thuiller, G. Lopez-Gonzalez, S. L. Lewis, M. Burkitt, O. L.
1156 Phillips, P. J. M. Maas, L. Y. T. Westra, H. Rainer, A. Q. Lobão, R. H. J.
1157 Erkens, P. V. A. Fine, D. C. Daly, G. V. Muñoz, I. Mesones, K. M. Cameron,
1158 K. J. Feeley, M. R. Silman, M. Dufrene, P. Legendre, J. Bunge, M. Fitzpatrick,
1159 J. Bunge, L. Woodard, D. Böhning, J. A. Foster, S. Connolly, H. K. Allen, J.-P.
1160 Z. Wang, B. G. Lindsay, A. Chao, R. K. Colwell, C.-W. Lin, N. J. Gotelli, C.
1161 X. Mao, R. K. Colwell, U. Brose, N. D. Martinez, R. J. Williams, I. J. Good, A.
1162 Chao, S.-M. Lee, I. Rocchetti, J. Bunge, D. Bohning, R. A. Fisher, A. S.
1163 Corbet, C. B. Williams, F. W. Preston, J.-P. Wang, P. V. A. Fine, and R. H.
1164 Ree, *Hyperdominance in the Amazonian Tree Flora*, *Science* **342**, 1243092
1165 (2013).
- 1166 [8] L. H. M. Cosme, J. Schietti, F. R. C. Costa, and R. S. Oliveira, *The Importance*
1167 *of Hydraulic Architecture to the Distribution Patterns of Trees in a Central*
1168 *Amazonian Forest*, *New Phytol.* **215**, 113 (2017).
- 1169 [9] S. E. Russo, S. J. Davies, D. A. King, and S. Tan, *Soil-Related Performance*
1170 *Variation and Distributions of Tree Species in a Bornean Rain Forest*, *J. Ecol.*
1171 **93**, 879 (2005).
- 1172 [10] K. J. Gaston, *Species-Range Size Distributions: Products of Speciation,*
1173 *Extinction and Transformation: Philosophical Transactions of the Royal*
1174 *Society of London Series B*, **353**, 219 (1998).
- 1175 [11] R. E. Ricklefs, H. Qian, and P. S. White, *The Region Effect on Mesoscale Plant*

- 1176 *Species Richness between Eastern Asia and Eastern North America*, **2**, 129
1177 (2004).
- 1178 [12] F. Munoz, B. R. Ramesh, and P. Couteron, *How Do Habitat Filtering and*
1179 *Niche Conservatism Affect Community Composition at Different Taxonomic*
1180 *Resolutions?*, *Ecology* **95**, 2179 (2014).
- 1181 [13] R. K. Colwell, C. Rahbek, and N. J. Gotelli, *The Mid-Domain Effect and*
1182 *Species Richness Patterns: What Have We Learned so Far?*, *Am. Nat.* **163**, E1
1183 (2004).
- 1184 [14] H. ter Steege, N. Pitman, D. Sabatier, H. Castellanos, P. Van der Hout, D. C.
1185 Daly, M. Silveira, O. Phillips, R. Thomas, J. V. A. N. Essen, H. Mogollon, and
1186 W. Morawetz, *A Spatial Model of Tree α -Diversity and Tree Density for the*
1187 *Amazon*, *Biodivers. Conserv.* **12**, 2255 (2003).
- 1188 [15] T. F. Rangel, N. R. Edwards, P. B. Holden, J. A. F. Diniz-Filho, W. D.
1189 Gosling, M. T. P. Coelho, F. A. S. Cassemiro, C. Rahbek, and R. K. Colwell,
1190 *Modeling the Ecology and Evolution of Biodiversity: Biogeographical Cradles,*
1191 *Museums, and Graves.* *Science*, Science (80-.). **in press**, (2018).
- 1192 [16] B. Shipley, *Measuring and Interpreting Trait-Based Selection versus Meta-*
1193 *Community Effects during Local Community Assembly*, *J. Veg. Sci.* **25**, 55
1194 (2014).
- 1195 [17] B. Shipley, C. E. T. Paine, and C. Baraloto, *Quantifying the Importance of*
1196 *Local Niche-Based and Stochastic Processes to Tropical Tree Community*
1197 *Assembly*, *Ecology* **93**, 760 (2012).
- 1198 [18] A. Colin Cameron and F. A. G. Windmeijer, *An R-Squared Measure of*
1199 *Goodness of Fit for Some Common Nonlinear Regression Models*, *J. Econom.*
1200 **77**, 329 (1997).
1201