
This is the **accepted version** of the article:

Bruelheide, Helge; Dengler, Jürgen; Purschke, Oliver; [et al.]. Global trait–environment relationships of plant communities. DOI 10.1038/s41559-018-0699-8

This version is available at <https://ddd.uab.cat/record/203561>

under the terms of the  **Free Access** license

1 **Title:**

2 **Global trait–environment relationships of plant communities**

3

4 **One Sentence Summary:** Trait composition of plant communities across the globe is
5 captured by two main dimensions and is probably shaped by environmental or biotic filtering,
6 but is only weakly related to global climate and soil gradients.

7 **Authors:**

8 Helge Bruelheide^{1,2,*}, Jürgen Dengler^{2,3,4}, Oliver Purschke^{1,2}, Jonathan Lenoir⁵, Borja
9 Jiménez-Alfaro^{6,1,2}, Stephan M. Hennekens⁷, Zoltán Botta-Dukát⁸, Milan Chytrý⁹, Richard
10 Field¹⁰, Florian Jansen¹¹, Jens Kattge^{2,12}, Valério D. Pillar¹³, Franziska Schrodtr^{10,12}, Miguel D.
11 Mahecha^{2,12}, Robert K. Peet¹⁴, Brody Sandel¹⁵, Peter van Bodegom¹⁶, Jan Altman¹⁷, Esteban
12 Alvarez Davila¹⁸, Mohammed A.S. Arfin Khan^{19,20}, Fabio Attorre²¹, Isabelle Aubin²²,
13 Christopher Baraloto²³, Jorcely G. Barroso²⁴, Marijn Bauters²⁵, Erwin Bergmeier²⁶, Idoia
14 Biurrun²⁷, Anne D. Bjorkman²⁸, Benjamin Blonder^{29,30}, Andraž Čarni^{31,32}, Luis Cayuela³³,
15 Tomáš Černý³⁴, J. Hans C. Cornelissen³⁵, Dylan Craven^{2,36}, Matteo Dainese³⁷, Géraldine
16 Derroire³⁸, Michele De Sanctis²¹, Sandra Díaz³⁹, Jiří Doležal¹⁷, William Farfan-Rios^{40,41}, Ted
17 R. Feldpausch⁴², Nicole J. Fenton⁴³, Eric Garnier⁴⁴, Greg R. Guerin⁴⁵, Alvaro G. Gutiérrez⁴⁶,
18 Sylvia Haider^{1,2}, Tarek Hattab⁴⁷, Greg Henry⁴⁸, Bruno Hérault^{49,50}, Pedro Higuchi⁵¹, Norbert
19 Hölzel⁵², Jürgen Homeier⁵³, Anke Jentsch²⁰, Norbert Jürgens⁵⁴, Zygmunt Kaçki⁵⁵, Dirk N.
20 Karger^{56,57}, Michael Kessler⁵⁶, Michael Kleyer⁵⁸, Ilona Knollová⁹, Andrey Y. Korolyuk⁵⁹,
21 Ingolf Kühn^{36,1,2}, Daniel C. Laughlin^{60,61}, Frederic Lens⁶², Jacqueline Loos⁶³, Frédérique
22 Louault⁶⁴, Mariyana I. Lyubenova⁶⁵, Yadvinder Malhi⁶⁶, Corrado Marcenò²⁷, Maurizio
23 Mencuccini^{67,68}, Jonas V. Müller⁶⁹, Jérôme Munzinger⁷⁰, Isla H. Myers-Smith⁷¹, David A.
24 Neill⁷², Ülo Niinemets⁷³, Kate H. Orwin⁷⁴, Wim A. Ozinga^{7,75}, Josep Penuelas^{68,73,76}, Aaron
25 Pérez-Haase^{77,78}, Petr Petřík¹⁷, Oliver L. Phillips⁷⁹, Meelis Pärtel⁸⁰, Peter B. Reich^{81,82},
26 Christine Römermann^{2,83}, Arthur V. Rodrigues⁸⁴, Francesco Maria Sabatini^{1,2}, Jordi
27 Sardans^{68,76}, Marco Schmidt⁸⁵, Gunnar Seidler¹, Javier Eduardo Silva Espejo⁸⁶, Marcos
28 Silveira⁸⁷, Anita Smyth⁴⁵, Maria Sporberr^{1,2}, Jens-Christian Svenning²⁸, Zhiyao Tang⁸⁸,
29 Raquel Thomas⁸⁹, Ioannis Tsiripidis⁹⁰, Kiril Vassilev⁹¹, Cyrille Violle⁴⁴, Risto Virtanen^{2,92,93},
30 Evan Weiher⁹⁴, Erik Welk^{1,2}, Karsten Wesche^{2,95,96}, Marten Winter², Christian Wirth^{2,12,97},
31 Ute Jandt^{1,2}

32 **Affiliations:**

33 ¹ Martin Luther University Halle-Wittenberg, Institute of Biology/Geobotany and Botanical
34 Garden, Am Kirchtor 1, 06108 Halle, Germany

35 ² German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Deutscher
36 Platz 5e, 04103 Leipzig, Germany

37 ³ Zurich University of Applied Sciences (ZHAW), Institute of Natural Resource Sciences
38 (IUNR), Research Group Vegetation Ecology, Grüentalstr. 14, Postfach, 8820 Wädenswil,
39 Switzerland

40 ⁴ University of Bayreuth, Bayreuth Center of Ecology and Environmental Research
41 BayCEER, Plant Ecology, Universitätsstr. 30, 95447 Bayreuth, Germany

42 ⁵ CNRS, Université de Picardie Jules Verne, UR «Ecologie et Dynamique des Systèmes
43 Anthropisés» (EDYSAN, UMR 7058 CNRS-UPJV), 1 rue des Louvels, 80037 Amiens Cedex
44 1, France

45 ⁶ Research Unit of Biodiversity (CSIC/UO/PA), University of Oviedo, Campus de Mieres, c/
46 Gonzalo Gutiérrez Quirós s/n 33600 Mieres, Spain

47 ⁷ Wageningen Environmental Research (Alterra), Team Vegetation, Forest and Landscape
48 Ecology, PO Box 47, 6700 AA Wageningen, The Netherlands

49 ⁸ MTA Centre for Ecological Research, GINOP Sustainable Ecosystems Group, 8237 Tihany,
50 Klebesberg Kuno u. 3, Hungary

51 ⁹ Masaryk University, Department of Botany and Zoology, Kotlářská 2, 611 37 Brno, Czech
52 Republic

53 ¹⁰ University of Nottingham, School of Geography, University Park, Nottingham, NG7 2RD,
54 United Kingdom

55 ¹¹ University of Rostock, Faculty for Agricultural and Environmental Sciences, Justus-von-
56 Liebig-Weg 6, 18059 Rostock, Germany

57 ¹² Max Planck Institute for Biogeochemistry, Hans-Knöll-Str. 10, 07745 Jena, Germany

58 ¹³ Universidade Federal do Rio Grande do Sul, Department of Ecology, Porto Alegre, RS,
59 91501-970, Brazil

60 ¹⁴ University of North Carolina at Chapel Hill, Department of Biology, Chapel Hill, NC
61 27599-3280, USA

62 ¹⁵ Santa Clara University, Department of Biology, Santa Clara CA 95053, USA

63 ¹⁶ Leiden University, Institute of Environmental Sciences, Department Conservation Biology,
64 Einsteinweg 2, 2333 CC Leiden, The Netherlands

65 ¹⁷ Institute of Botany of the Czech Academy of Sciences, Zámek 1, 252 43 Průhonice, Czech
66 Republic

67 ¹⁸ Escuela de Ciencias Agropecuarias y Ambientales – ECAPMA, Universidad Nacional
68 Abierta y a Distancia –UNAD, Sede José Celestino Mutis, Calle 14S #14, Bogotá

69 ¹⁹ Shahjalal University of Science and Technology, Department of Forestry and
70 Environmental Science, Sylhet, 3114, Bangladesh

71 ²⁰ University of Bayreuth, Bayreuth Center of Ecology and Environmental Research
72 BayCEER, Department of Disturbance Ecology, Universitätsstr. 30, 95447 Bayreuth,
73 Germany
74 ²¹ Sapienza University of Rome, Department of Environmental Biology, P.le Aldo Moro 5,
75 00185 Rome, Italy
76 ²² Great Lakes Forestry Centre, Canadian Forest Service, Natural Resources Canada, 1219
77 Queen St. East, Sault Ste Marie, ON, P6A 2E5, Canada
78 ²³ Florida International University, Department of Biological Sciences, International Center
79 for Tropical Botany (ICTB), 11200 SW 8th Street, OE 243 Miami, FL 33199, USA
80 ²⁴ Universidade Federal do Acre, Campus de Cruzeiro do Sul, Acre, Brazil.
81 ²⁵ Ghent University, Faculty of Bioscience Engineering, Department of Applied Analytical
82 and Physical Chemistry, ISOFYS, Coupure Links 653, 9000 Gent, Belgium
83 ²⁶ University of Göttingen, Albrecht von Haller Institute of Plant Sciences, Vegetation
84 Analysis & Plant Diversity, Untere Karspüle 2, 37073 Göttingen, Germany
85 ²⁷ University of the Basque Country UPV/EHU, Apdo. 644, 48080 Bilbao, Spain
86 ²⁸ Aarhus University, Department of Bioscience, Section for Ecoinformatics & Biodiversity,
87 Ny Munkegade 114, 8000 Aarhus C, Denmark
88 ²⁹ University of Oxford, Environmental Change Institute, School of Geography and the
89 Environment, South Parks Road, Oxford OX1 3QY, United Kingdom
90 ³⁰ Rocky Mountain Biological Laboratory, PO Box 519, Crested Butte, Colorado, 81224 USA
91 ³¹ Scientific Research Center of the Slovenian Academy of Sciences and Arts, Institute of
92 Biology, Novi trg 2, SI 1001 Ljubljana p. box 306, Slovenia
93 ³² University of Nova Gorica, 5000 Nova Gorica, Slovenia
94 ³³ Universidad Rey Juan Carlos, Department of Biology, Geology, Physics and Inorganic
95 Chemistry, c/ Tulipán s/n, 28933 Móstoles, Madrid, Spain
96 ³⁴ Czech University of Life Sciences, Faculty of Forestry and Wood Science, Department of
97 Forest Ecology, Kamýcká 1176, 16521, Praha 6 – Suchdol, Czech Republic
98 ³⁵ Vrije Universiteit Amsterdam, Faculty of Science, Department of Ecological Science, De
99 Boelelaan 1085, 1081 HV Amsterdam, The Netherlands
100 ³⁶ Helmholtz Centre for Environmental Research - UFZ, Dept. Community Ecology,
101 Theodor-Lieser-Str. 4, 06120 Halle, Germany
102 ³⁷ University of Würzburg, Department of Animal Ecology and Tropical Biology, Am
103 Hubland, 97074 Würzburg, Germany
104 ³⁸ Cirad, UMR EcoFoG, Campus Agronomique, 97310 Kourou, French Guiana

105 ³⁹ Instituto Multidisciplinario de Biología Vegetal (IMBIV), CONICET and FCEfyN,
106 Universidad Nacional de Córdoba, Casilla de Correo 495, 5000 Córdoba, Argentina.

107 ⁴⁰ Wake Forest University, Department of Biology, Winston Salem, North Carolina, USA

108 ⁴¹ Universidad Nacional de San Antonio Abad del Cusco, Herbario Vargas (CUZ), Cusco,
109 Peru

110 ⁴² University of Exeter, College of Life and Environmental Sciences, Geography, Exeter, EX4
111 4RJ, United Kingdom.

112 ⁴³ Université du Québec en Abitibi-Témiscamingue, Institut de recherche sur les forêts, 445
113 Boul. de l'Université, Rouyn-Noranda, Qc J9X 4E5, Canada

114 ⁴⁴ CNRS - Université de Montpellier - Université Paul-Valéry Montpellier - EPHE, Centre
115 d'Ecologie Fonctionnelle et Evolutive (UMR5175), 34293 Montpellier Cedex 5, France

116 ⁴⁵ University of Adelaide, Terrestrial Ecosystem Research Network, School of Biological
117 Sciences, Adelaide, South Australia, 5005 Australia

118 ⁴⁶ Universidad de Chile, Facultad de Ciencias Agronómicas, Departamento de Ciencias
119 Ambientales y Recursos Naturales Renovables, Av. Santa Rosa 11315, La Pintana 8820808,
120 Santiago, Chile

121 ⁴⁷ IFREMER (Institut Français de Recherche pour l'Exploitation de la MER) UMR 248
122 MARBEC (CNRS, IFREMER, IRD, UM), 34203 Sète cedex, France

123 ⁴⁸ University of British Columbia, The Department of Geography, 1984 West Mall,
124 Vancouver, BC V6T 1Z2, Canada

125 ⁴⁹ INPHB (Institut National Polytechnique Félix Houphouët-Boigny), BP 1093,
126 Yamoussoukro, Ivory Coast

127 ⁵⁰ Cirad, University Montpellier, UR Forests & Societies, Montpellier, France

128 ⁵¹ Universidade do Estado de Santa Catarina (UDESC), Departamento de Engenharia
129 Florestal, Av Luiz de Camões, 2090 - Conta Dinheiro, Lages – SC, 88.520 – 000, Brazil

130 ⁵² University of Münster, Institute of Landscape Ecology, Heisenbergstr. 2, 48149 Münster,
131 Germany

132 ⁵³ University of Göttingen, Plant Ecology and Ecosystems Research, Untere Karspüle 2,
133 37073 Göttingen, Germany

134 ⁵⁴ University of Hamburg, Biodiversity, Biocenter Klein Flottbek and Botanical Garden,
135 Ohnhorststr. 18, 22609 Hamburg, Germany

136 ⁵⁵ University of Wrocław, Institute of Environmental Biology, Department of Vegetation
137 Ecology, Przybyszewskiego 63, 51-148 Wrocław, Poland

138 ⁵⁶ University of Zurich, Department of Systematic and Evolutionary Botany, Zollikerstrasse
139 107, 8008 Zurich, Switzerland

140 ⁵⁷ Swiss Federal Research Institute WSL, Zürcherstrasse 111, 8903 Birmensdorf, Switzerland.

141 ⁵⁸ University of Oldenburg, Institute of Biology and Environmental Sciences, Landscape
142 Ecology Group, Carl-von-Ossietzky Strasse 9-11, 26111 Oldenburg, Germany

143 ⁵⁹ Central Siberian Botanical Garden SB RAS, Zolotodolinskaya str. 101, Novosibirsk,
144 630090, Russia

145 ⁶⁰ University of Waikato, Environmental Research Institute, School of Science, Private Bag
146 3105, Hamilton 3240, New Zealand

147 ⁶¹ University of Wyoming, Department of Botany, Laramie, Wyoming, USA

148 ⁶² Leiden University, Naturalis Biodiversity Center, P.O. Box 9517, 2300RA Leiden, The
149 Netherlands

150 ⁶³ Agroecology, University of Göttingen, Grisebachstrasse 6, 37077 Göttingen, Germany

151 ⁶⁴ INRA, VetAgro Sup, UMR Ecosystème Prairial, 63000 Clermont-Ferrand, France

152 ⁶⁵ University of Sofia, Faculty of Biology, Department of Ecology and Environmental
153 Protection, 1164 Sofia, 8 Dragan Tsankov Av., Bulgaria

154 ⁶⁶ University of Oxford, Environmental Change Institute, School of Geography and the
155 Environment, South Parks Road, Oxford, OX1 3QY, United Kingdom

156 ⁶⁷ ICREA Pg. Lluís Companys 23, 08010 Barcelona, Spain

157 ⁶⁸ CREAM, Cerdanyola del Vallès, 08193 Barcelona, Catalonia, Spain

158 ⁶⁹ Royal Botanic Gardens Kew, Millennium Seed Bank, Conservation Science, Wakehurst
159 Place, Ardingly, RH17 6TN, United Kingdom

160 ⁷⁰ AMAP, IRD, CNRS, INRA, Université Montpellier, 34000 Montpellier, France

161 ⁷¹ University of Edinburgh, School of GeoSciences, Edinburgh EH9 3FF, United Kingdom

162 ⁷² Universidad Estatal Amazónica, Conservación y Manejo de Vida Silvestre, Paso lateral km
163 2½ via a Napo Puyo, Pastaza, Ecuador

164 ⁷³ Estonian University of Life Science, Department of Crop Science and Plant Biology,
165 Kreutzwaldi 1, 51014, Tartu, Estonia

166 ⁷⁴ Landcare Research, PO Box 69040, Lincoln 7640, New Zealand

167 ⁷⁵ Radboud University Nijmegen, Institute for Water and Wetland Research, 6500 GL
168 Nijmegen, The Netherlands

169 ⁷⁶ CSIC, Global Ecology Unit, CREAM-CEAB-UAB, Cerdanyola del Vallès, 08193
170 Barcelona, Catalonia, Spain

171 ⁷⁷ University of Barcelona, Faculty of Biology, Department of Evolutionary Biology, Ecology
172 and Environmental Sciences, Av. Diagonal 643, 08028, Barcelona, Spain
173 ⁷⁸ Center for Advanced Studies of Blanes, Spanish Research Council
174 (CEAB-CSIC), 17300 Blanes, Spain
175 ⁷⁹ University of Leeds, School of Geography, Leeds LS2 9JT, United Kingdom
176 ⁸⁰ University of Tartu, Ülikooli 18, 50090 Tartu, Estonia
177 ⁸¹ University of Minnesota, Department of Forest Recourses, 220F Green Hall,
178 1530 Cleveland Avenue North, St. Paul, MN 55108, USA
179 ⁸² Western Sydney University, Hawkesbury Institute for the Environment, New South Wales
180 2751, Australia
181 ⁸³ Friedrich Schiller University Jena, Institute of Systematic Botany, Philosophenweg 16,
182 07743 Jena, Germany
183 ⁸⁴ Universidade Regional de Blumenau, Departamento de Engenharia Florestal, Rua São
184 Paulo, 3250, 89030-000 Blumenau-Santa Catarina, Brazil
185 ⁸⁵ Senckenberg Biodiversity and Climate Research Centre (BiK-F), Data and Modelling
186 Centre, Senckenberganlage 25, 60325 Frankfurt am Main, Germany
187 ⁸⁶ University of La Serena , Department of Biology, La Serena, Chile
188 ⁸⁷ Universidade Federal do Acre, Museu Universitário / Centro de Ciências Biológicas e da
189 Natureza / Laboratório de Botânica e Ecologia Vegetal, BR 364, Km 04 - Distrito Industrial -
190 69915-559 - Rio Branco-AC, Brazil
191 ⁸⁸ Peking University, College of Urban and Environmental Sciences, Yiheyuan Rd. 5, 100871,
192 Beijing, China
193 ⁸⁹ Iwokrama International Centre for Rain Forest Conservation and Development, 77 High
194 Street, Kingston, Georgetown, Guyana
195 ⁹⁰ Aristotle University of Thessaloniki, School of Biology, Department of Botany, 54124
196 Greece
197 ⁹¹ Bulgarian Academy of Sciences, Institute of Biodiversity and Ecosystem Research, 23
198 Acad. G. Bonchev Str., 1113 Sofia, Bulgaria
199 ⁹² Helmholtz Center for Environmental Research – UFZ, Department of Physiological
200 Diversity, Permoserstraße 15, Leipzig 04318, Germany
201 ⁹³ University of Oulu, Department of Ecology & Genetics, PO Box 3000, FI-90014, Finland
202 ⁹⁴ University of Wisconsin - Eau Claire, Department of Biology, Eau Claire, WI 54702-4004,
203 USA

204 ⁹⁵ Senckenberg Museum of Natural History Görlitz, P.O. Box 300154, 02806 Görlitz,
205 Germany

206 ⁹⁶ TU Dresden, International Institute (IHI) Zittau, Markt 23, 02763 Zittau, Germany

207 ⁹⁷ University of Leipzig, Systematic Botany and Functional Biodiversity, Johannisallee 21–
208 23, 04103 Leipzig, Germany

209

210 * Corresponding author

211

212

213 **Abstract:**

214 Plant functional traits directly affect ecosystem functions. At the species level, trait
215 combinations depend on trade-offs representing different ecological strategies, but at the
216 community level trait combinations are expected to be decoupled from these trade-offs
217 because different strategies can facilitate co-existence within communities. A key remaining
218 question is to what extent community-level trait composition is globally filtered and how well
219 it is related to global vs. local environmental drivers. Here, we perform a global, plot-level
220 analysis of trait-environment relationships, using a database with more than 1.1 million
221 vegetation plots and 26,632 plant species with trait information. Although we found a strong
222 filtering of 17 functional traits, similar climate and soil conditions support communities
223 differing greatly in mean trait values. The two main community trait axes which capture half
224 of the global trait variation (plant stature and resource acquisitiveness) reflect the trade-offs at
225 the species level but are weakly associated with climate and soil conditions at the global scale.
226 Similarly, within-plot trait variation does not vary systematically with macro-environment.
227 Our results indicate that, at fine spatial grain, macro-environmental drivers are much less
228 important for functional trait composition than has been assumed from floristic analyses
229 restricted to co-occurrence in large grid cells. Instead, trait combinations seem to be
230 predominantly filtered by local-scale factors such as disturbance, fine-scale soil conditions,
231 niche partitioning or biotic interactions.

232

233 **Introduction**

234 How climate drives the functional characteristics of vegetation across the globe has been a
235 key question in ecological research for more than a century¹. While functional information is
236 available for a large portion of the global pool of plant species, we do not know how
237 functional traits of the different species that co-occur in a community are combined, which is
238 what determines their joint effect on ecosystems²⁻⁴. At the species level, Díaz et al.⁵
239 demonstrated that 74% of the global spectrum of six key plant traits determining plant fitness
240 in terms of survival, growth and reproduction can be accounted for by two principal
241 components (PCs). They showed that the functional space occupied by vascular plant species
242 is strongly constrained by trade-offs between traits and converges on a small set of successful
243 trait combinations, confirming previous findings⁶⁻⁹. While these constraints describe
244 evolutionarily viable ecological strategies for vascular plant species globally, they provide
245 only limited insight into trait composition within communities. There are many reasons why
246 trait composition within communities would produce very different patterns, and indeed much
247 theory predicts this¹⁰⁻¹¹. However, it is still unknown to what extent community-level trait
248 composition depends on local factors (microclimate, fine-scale soil properties, disturbance
249 regime¹⁰, successional dynamics²) and regional to global environmental drivers
250 (macroclimate^{6,12-13}, coarse-scale soil properties^{3,14}). As ecosystem functions and services are
251 ultimately dependent on the traits of the species composing ecological communities,
252 exploring community trait composition at the global scale can advance our understanding of
253 how climate change and other anthropogenic drivers affect ecosystem functioning.

254 So far, studies relating trait composition to the environment at continental to global extents
255 have been restricted to coarse-grained species occurrence data (e.g. presence in 1° grid cells¹⁵⁻
256 ¹⁷). Such data capture neither biotic interactions (co-occurrence in large grid cells does not
257 indicate local co-existence), nor local variation in environmental filters (e.g. variation in soil,
258 topography or disturbance regime within grid cells). In contrast, functional composition of
259 ecological communities sampled at fine-grained vegetation plots – with areas of few to a few
260 hundred square meters – is the direct outcome of the interaction between both local and large-
261 scale factors. Here, we present a global analysis of plot-level trait composition. We combined
262 the ‘sPlot’ database, a new global initiative incorporating more than 1.1 million vegetation
263 plots from over 100 databases (mainly forests and grasslands; see Methods), with 30 large-
264 scale environmental variables and 18 key plant functional traits derived from TRY, a global
265 plant-trait database (see Methods, Table 2). We selected these 18 traits because they affect
266 different key ecosystem processes and are expected to respond to macroclimatic drivers
267 (Table 1). In addition, they were sufficiently measured across all species globally to allow for
268 imputation of missing values (see Methods). All analyses were confined to vascular plant
269 species and included all vegetation layers in a community, from the canopy to the herb layer
270 (see Methods).

271 We used this unprecedented fine-resolution dataset to test the hypothesis (Hypothesis 1) that
272 plant communities show evidence of environmental or biotic filtering at the global scale,
273 making use of the observed variation of plot-level trait means and means of within-plot trait
274 variation across communities. Ecological theory suggests that community-level convergence
275 could be interpreted as the result of filtering processes, including environmental filtering and
276 biotic interactions. Globally, temperature and precipitation drive the differences in vegetation
277 between biomes, suggesting strong environmental filtering^{3,11} that constrains the number of
278 successful trait combinations and leads to community-level trait convergence. Similarly,
279 biotic interactions may eliminate excessively divergent trait combinations^{18,19}. However,
280 alternative functional trait combinations may confer equal fitness in the same environment¹⁰.
281 If plant communities show a global variation of plot-level trait means higher than expected by
282 chance, and a lower than expected within-plot trait variation (see Figure 1), this would
283 support the view that environmental or biotic filtering are dominant structuring processes of
284 community trait composition at the global scale. A consequence of strong community-level
285 trait convergence, and thus low variation within plots with species trait values centred around
286 the mean, would be that plot-level means will be similar to the trait values of the species in
287 that plot. Hence, community mean trait values should then mirror the trait values of individual
288 species⁵.

289 While Hypothesis 1 addresses the degree of filtering, it does not make a statement on the
290 attribution of driving factors. The main drivers should correlate strongly (though not
291 necessarily linearly²⁰) with plot-level trait means and within-plot trait variance. Identifying
292 these drivers has the potential to fundamentally improve our understanding of global trait-
293 environment relationships. We tested the hypothesis (Hypothesis 2) that there are strong
294 correlations between global environmental drivers such as macroclimate and coarse-scale soil
295 properties and both plot-level trait means and within-plot trait variances^{3,6,12-17,20-24} (see Table
296 1 for expected relationships and Supplementary Table 2 for variables used). Such evidence,

297 although correlative, may contribute to the formulation of novel hypotheses to explain global
298 plant trait patterns.

299

300 **Results and Discussion**

301 Consistent with Hypothesis 1 and as illustrated in Figure 1, global variation in plot-level trait
302 means was much higher than expected by chance: all traits had positive standardized effect
303 sizes (SEs), which were significantly > 0 for 17 out of 18 traits based on gap-filled data
304 (mean SES = 8.06 standard deviations (SD), Table 2). This suggests that environmental or
305 biotic filtering is a dominant force of community trait composition globally. Also as predicted
306 by Hypothesis 1, within-plot trait variance was typically lower than expected by chance
307 (mean SES = -1.76 SD, significantly < 0 for ten traits but significantly > 0 for three traits;
308 Table 2). Thus, trait variation within communities may also be constrained by filtering.

309 Trait correlations at the community level were relatively well captured by the first two axes of
310 a Principal Component Analysis (PCA) for both plot-level trait means and within-plot trait
311 variances (Figures 1 and 2). The dominant axes were determined by those traits with the
312 highest absolute SEs of plot-level trait mean trait values (Table 2, mean of CWMs). The
313 PCA of plot-level trait means (Fig. 2) reflects two main functional continua on which
314 community trait values converge: one from short-stature, small-seeded communities such as
315 grasslands or herbaceous vegetation to tall-stature communities with large, heavy diaspores
316 such as forests (the size spectrum), and the other from communities with resource-acquisitive
317 to those with resource-conservative leaves (i.e. the leaf economics spectrum)⁷. The high
318 similarity between this PCA and the one at the species level by Díaz et al.⁵ is striking: here at
319 the community level, based on 1.1 million plots, the same functional continua emerged as at
320 the species level, based on 2,214 species. While the trade-offs between different traits at the
321 species level can be understood from a physiological and evolutionary perspective, finding
322 similar trade-offs between traits at the community level was unexpected, as species with
323 opposing trait values can co-exist in the same community. In combination with our finding of
324 strong trait convergence, these results reveal a strong parallel of present-day community
325 assembly to individual species' evolutionary histories.

326 Surprisingly, we found only limited support for Hypothesis 2. Community-level trait
327 composition was poorly captured by global climate and soil variables. None of the 30
328 environmental variables accounted individually for more than 10% of the variance in the traits
329 defining the main dimensions in Fig. 2 (Supplementary Fig. 2). The coefficients of
330 determination were not improved when testing for non-linear relationships (see Methods).
331 Using all 30 environmental variables simultaneously as predictors only accounted for 10.8%
332 or 14.0% of the overall variation in plot-level trait means (cumulative variance, respectively,
333 of the first two or all 18 constrained axes in a Redundancy Analysis). Overall, our results
334 show that similar global-scale climate and soil conditions can support communities that differ
335 markedly in mean trait values and that different climates can support communities with rather
336 similar mean trait values.

337 The ordination of within-plot variance of the different traits (Fig. 3) revealed two main
338 continua. Variances of plant height and diaspore mass varied largely independently of
339 variances of traits representing the leaf economics spectrum. This suggests that short and tall
340 species can be assembled together in the same community independently from combining
341 species with acquisitive leaves with species with conservative leaves. Global climate and soil
342 variables accounted for even less variation on the first two PCA axes in within-plot trait
343 variances than on the first two PCA axes in plot-level trait means. Only two environmental
344 variables had $r^2 > 3\%$ (Supplementary Fig. 3), whether allowing for non-linear relationships
345 (see Methods) or not, and overall, macro-environment accounted for only 3.6% or 5.0% of the
346 variation (cumulative variance, respectively, of the first two or all 18 constrained axes).
347 Removing species richness effects from within-plot trait variances did not increase the
348 amount of variation explained by the environment (see Methods).

349 The findings of our study contrast strongly with studies where the variation in traits between
350 species was calculated at the level of the species pool in large grid cells^{15,16}, suggesting that
351 plot-level and grid cell-level trait composition are driven by different factors²¹. Plot-level trait
352 means and variances may both be predominantly driven by local environmental factors, such
353 as topography (e.g. north- vs. south-facing slopes), local soil characteristics (e.g. soil depth
354 and nutrient supply)^{3,14,24,25}, disturbance regime (including land use²⁶ and successional
355 status^{2,27}) or biotic interactions^{18-19,28}, while broad-scale climate and soil conditions may only
356 become relevant for the whole species pool in large grid cells. Such differences emphasize the
357 importance of local environment in affecting the communities' trait composition and should
358 be taken into account when interpreting the effect of environmental drivers in functional trait
359 diversity using data on either floristic pools or ecological communities.

360 We note that the strongest community-level correlations with environment were found for
361 traits not linked to the leaf economics spectrum. Mean stem specific density increased with
362 potential evapotranspiration (PET, $r^2=15.6\%$; Fig. 4a, b), reflecting the need to produce
363 denser wood with increasing evaporative demand. Leaf N:P ratio increased with growing-
364 season warmth (growing degree days above 5°C, GDD5, $r^2=11.5\%$; Fig. 4d), indicating strong
365 phosphorus limitation²⁹ in most plots in the tropics and subtropics (Fig. 4c, d). This pattern
366 was not brought about by a parallel increase in the presence of legumes, which tend to have
367 relatively high N:P ratios; excluding all species of Fabaceae resulted in a very similar
368 relationship with GDD5 ($r^2=10.0\%$). The global N:P pattern is consistent with results based
369 on traits of single species related to mean annual temperature³⁰. We assume that the main
370 underlying mechanism is the high soil weathering rate at high temperatures and humidity,
371 which in the tropics and subtropics was not reset by Pleistocene glaciation. Thus, phosphorus
372 limitation may weaken the relationships between productivity-related traits and macroclimate
373 (Supplementary Fig. 2). For example, specific leaf area may be low as consequence of low
374 nutrient availability^{3,14,24-25} in favourable climates as well as be low as consequence of low
375 temperature and precipitation under favourable nutrient supply. Overall, our findings are
376 relevant in improving Dynamic Global Vegetation Models (DGVMs), which so far have used
377 trait information only from a few calibration plots²². The sPlot database provides much-
378 needed empirical data on the community trait pool in DGVMs³¹ and identifies traits that

379 should be considered when predicting ecosystem functions from vegetation, such as stem
380 specific density and leaf N:P ratio.

381 Our results were surprisingly robust both to the selection of trait data, when comparing
382 different plant formations and when explicitly accounting for the uneven distribution of plots.
383 Using the original trait values measured for the species from the TRY database for the six
384 traits used by Díaz et al.⁵ (see Methods), resulted in the same two main functional continua
385 and an overall highly similar ordination pattern (Supplementary Fig. 4) compared to using
386 gap-filled data for 18 traits (Fig. 2). Community-level trait composition was also similarly
387 poorly captured by global climate and soil variables. Single regressions of CWMs with all
388 environmental variables revealed very similar patterns to those based on gap-filled traits
389 (Supplementary Fig. 5). Similarly, subjecting the CWMs based on six original traits to a
390 Redundancy Analysis with all 30 environmental variables accounted only for 20.6% or 21.8%
391 of the overall variation in CWMs (cumulative variance of the first two or all six constrained
392 axes, respectively, Supplementary Fig. 4). These results clearly demonstrate that the
393 imputation of missing trait values did not result in spurious artefacts which may have
394 obscured community trait-environment relationships.

395 We also assessed whether the observed trait-environment relationships hold for forests and
396 non-forest vegetation independently (see Methods). Both subsets confirmed the overall
397 patterns in trait means (Supplementary Figs. 3-6). The variance in plot-level trait means
398 explained by large-scale climate and soil variables was higher for forest than non-forest plots,
399 probably because forests belong to a well-defined and rather resource-conservative formation,
400 whereas non-forest plots encompass a heterogeneous mixture of different vegetation types,
401 ranging from alpine meadows to semi-deserts, and tend to depend more on disturbance and
402 management, which can strongly affect trait-environment relationships of communities²¹.
403 Finally, to test whether our findings depended on the uneven distribution of plots among the
404 world's different climates and soils, we repeated the analyses in 100 subsets of ~100,000 plots
405 resampled in the global climate space (Supplementary Figs. 7-8). The analyses of the
406 resampled datasets revealed the same patterns and confirmed the impact of PET and GDD5
407 on stem specific density and leaf N:P ratio, respectively. The correlations between trait means
408 and environmental variables were, however, stronger in the resampled subsets, possibly
409 because the resampling procedure reduced the overrepresentation of the temperate-zone areas
410 with intermediate climatic values.

411 Our findings have important implications for understanding and predicting plant community
412 trait assembly. First, worldwide trait variation of plant communities is captured by a few main
413 dimensions of variation, which are surprisingly similar to those reported by species-based
414 studies^{5,7-9}, suggesting that the drivers of past trait evolution, which resulted in the present-day
415 species-level trait spectra⁵, are also reflected in the composition of today's plant communities.
416 If species-level trade-offs indeed constrain community assembly, then the present-day
417 contrasts in trait composition of terrestrial plant communities should also have existed in the
418 past and will probably remain, even for novel communities, in the future. Most species in our
419 present-day communities evolved under very variable filtering conditions across the globe,
420 with respect to temperature and precipitation regimes. Therefore, it can be assumed that future
421 filtering conditions will result in novel communities that follow the same functional continua

422 from short-stature, small-seeded communities to tall-stature communities with large, heavy
423 diaspores and from communities with resource-acquisitive to those with resource-
424 conservative leaves. Second, the main plot-level vegetation trait continua cannot easily be
425 captured by coarse-resolution environmental variables²¹. This brings into question both the
426 use of simple large-scale climate relationships to predict the leaf economics spectra of global
427 vegetation^{13,15-16,22} and attempts to derive net primary productivity and global carbon and
428 water budgets from global climate, even when employing powerful trait-based vegetation
429 models³¹. The finding that within-plot trait variances were only very weakly related to global
430 climate or soil variables points to the importance of i) local-scale climate or soil variables, ii)
431 disturbance regimes or iii) biotic interactions for the degree of local trait dispersion¹¹. Finally,
432 both our findings on the limited role of large-scale climate in explaining trait patterns and on
433 the prevalence of phosphorus limitation in most plots in the tropics and subtropics call for
434 including local variables when predicting community trait patterns. Even under similar
435 macro-environmental conditions, communities can vary greatly in trait means and variances,
436 consistent with high local variation in species' trait values^{3,6-7}. Future research on functional
437 response of communities to changing climate should incorporate the effect of local
438 environmental conditions²⁴⁻²⁶ and biotic interactions¹⁸⁻¹⁹ for building reliable predictions of
439 vegetation dynamics.

440

441 **Material and Methods**

442 **Vegetation Data.** The sPlot 2.1 vegetation database contains 1,121,244 plots with 23,586,216
443 species × plot observations, i.e. records of a species in a plot
444 (https://www.idiv.de/en/sdiv/working_groups/wg_pool/splot.html). This database aims at
445 compiling plot-based vegetation data from all vegetation types worldwide, but with a
446 particular focus on forest and grassland vegetation. Although the initial aim of sPlot was to
447 achieve global coverage, the plots are very unevenly distributed with most data coming from
448 Europe, North America and Australia and an overrepresentation of temperate vegetation types
449 (Supplementary Fig. 1).

450 For most plots (97.2%) information on the single species' relative contribution to the sum of
451 plants in the plot was available, expressed as cover, basal area, individual count, importance
452 value or per cent frequency in subplots. For the other 2.8% (31,461 plots), for which only
453 presence/absence (p/a) was available, we assigned equal relative abundance to the species
454 (1/species richness). For plots with a mix of cover and p/a information (mostly forest plots,
455 where herb layer information had been added on a p/a basis; 8,524 plots), relative abundance
456 was calculated by assigning the smallest cover value that occurred in a particular plot to all
457 species with only p/a information in that plot. In most cases (98.4%), plot records in sPlot
458 include full species lists of vascular plants. Bryophytes and lichens were additionally
459 identified in 14% and 7% of plots, respectively. After removing plots without geographic
460 coordinates and all observations on bryophytes and lichens, the database contained
461 22,195,966 observations on the relative abundance of vascular plant species in a total of
462 1,117,369 plots. The temporal extent of the data spans from 1885 to 2015, but >95% of
463 vegetation plots were recorded later than 1980. Plot size was reported in 65.4% of plots.

464 While forest plots had plot sizes $\geq 100 \text{ m}^2$, and in most cases $\leq 1,000 \text{ m}^2$, non-forest plots
465 typically ranged from 5 to 100 m^2 .

466 **Taxonomy.** To standardize the nomenclature of species within and between sPlot and TRY
467 (see below), we constructed a taxonomic backbone of the 121,861 names contained in the two
468 databases. Prior to name matching, we ran a series of string manipulation routines in R, to
469 remove special characters and numbers, as well as standardized abbreviations in names.
470 Taxon names were parsed and resolved using Taxonomic Name Resolution Service version
471 4.0 (TNRS³²; <http://tnrs.iplantcollaborative.org>; accessed 20 Sep 2015), selecting the best
472 match across the five following sources: i) The Plant List (version 1.1;
473 <http://www.theplantlist.org/>; Accessed 19 Aug 2015), ii) Global Compositae Checklist (GCC,
474 <http://compositae.landcareresearch.co.nz/Default.aspx>; accessed 21 Aug 2015), iii)
475 International Legume Database and Information Service (ILDIS,
476 <http://www.ildis.org/LegumeWeb>; accessed 21 Aug 2015), iv) Tropicos
477 (<http://www.tropicos.org/>; accessed 19 Dec 2014), and v) [USDA Plants Database](http://usda.gov/wps/portal/usda/usdahome)
478 (<http://usda.gov/wps/portal/usda/usdahome>; accessed 17 Jan 2015). We allowed for partial
479 matching to the next higher taxonomic rank (genus or family) in cases where full taxon names
480 could not be found. All names matched or converted from a synonym by TNRS were
481 considered accepted taxon names. In cases when no exact match was found (e.g. when
482 alternative spelling corrections were reported), names with probabilities of $\geq 95\%$ or higher
483 were accepted and those with $< 95\%$ were examined individually. Remaining non-matching
484 names were resolved based on the National Center for Biotechnology Information's
485 Taxonomy database (NCBI, <http://www.ncbi.nlm.nih.gov/>; accessed 25 Oct 2011) within
486 TNRS, or sequentially compared directly against The Plant List and Tropicos (accessed
487 September 2015). Names that could not be resolved against any of these lists were left as
488 blanks in the final standardized name field. This resulted in a total of 86,760 resolved names,
489 corresponding to 664 families, occurring in sPlot or TRY or both. Classification into families
490 was carried out according to APGIII³³, and was used to identify non-vascular plant species
491 ($\sim 5.1\%$ of the taxon names) which were excluded from the subsequent statistical analysis.

492 **Trait Data.** Data for 18 traits that are ecologically relevant (Table 1) and sufficiently covered
493 across species³⁴ were requested from TRY³⁵ (version 3.0) on the 10th August, 2016. We
494 applied gap-filling with Bayesian Hierarchical Probabilistic Matrix Factorization
495 (BHPMF^{34,36-37}). We used the prediction uncertainties provided by BHPMF for each
496 imputation to assess the quality of gap-filling and removed all imputations with a coefficient
497 of variation > 1 ³⁷. We obtained 18 gap-filled traits for 26,632 out of a total of 58,065 taxa in
498 sPlot, which corresponds to 45.9% of all species but to 88.7% of all species \times plot
499 combinations. Trait coverage of the most frequent species was 77.2% and 96.2% for taxa that
500 occurred in more than 100 or 1,000 plots, respectively. The gap-filled trait data comprised
501 observed and imputed values on 632,938 individual plants, which we \log_e transformed and
502 aggregated by taxon. For those taxa that were recorded at the genus level only, we calculated
503 genus means. Out of 22,195,966 records of vascular plant species with geographic reference,
504 21,172,989 ($=95.4\%$) refer to taxa for which we had gap-filled trait values. This resulted in
505 1,115,785 and 1,099,463 plots for which we had at least one taxon or two taxa with a trait

506 value (99.5% and 98.1%, respectively, of all 1,121,244 plots), and for which trait means and
507 variances could be calculated.

508 As some mean values of traits in TRY were based on a very small number of replicates per
509 species, which results in uncertainty in trait mean and variance calculations³⁸, we tested to
510 which degree the trait patterns in the dataset might be caused by a potential removal of trait
511 variation by imputation of trait values and additionally carried out all analyses using the
512 original trait data on the same 632,938 individual plants instead of gap-filled data
513 (Supplementary Table 1). The degree of trait coverage of species ranged between 7.0% and
514 58.0% for leaf fresh mass and plant height, respectively. Across all species, mean coverage of
515 species with original trait values was 21.8%, as compared to 45.9% for gap-filled trait data.
516 Linking these trait values to the species occurrence data resulted in a coverage of species ×
517 plot observations with trait values between 7.6% and 96.6% for conduit element length and
518 plant height, respectively (Supplementary Table 1), with a mean of 60.7% as compared to
519 88.7% for those based on gap-filled traits. Using these original trait values to calculate
520 community-weighted mean (CWM) trait values (see below) resulted in a plot coverage of trait
521 values between 48.2% and 100% for conduit element length and SLA, respectively. Across all
522 plots, mean coverage of plots with original trait values was 89.3%, as compared to 100% for
523 gap-filled trait data (Supplementary Table 1).

524 We are aware that using species mean values for traits excludes the possibility to account for
525 intraspecific variance, which can also strongly respond to the environment³⁹. Thus, using one
526 single value for a species is a source of error in calculating trait means and variances.

527

528 **Environmental Data.** We compiled 30 environmental variables (Supplementary Table 2).
529 Macroclimate variables were extracted from CHELSA⁴⁰⁻⁴¹, V1.1 (Climatologies at High
530 Resolution for the Earth's Land Surface Areas, www.chelsa-climate.org). CHELSA provides
531 19 bioclimatic variables equivalent to those used in WorldClim (www.worldclim.org) at a
532 resolution of 30 arc sec (~ 1 km at the equator), averaging global climatic data from the
533 period 1979–2013 and using a quasi-mechanistic statistical downscaling of the ERA-Interim
534 reanalysis⁴².

535 Variables reflecting growing-season warmth were growing degree days above 1°C (GDD1)
536 and 5°C (GDD5), calculated from CHELSA data⁴³. We also compiled an index of aridity
537 (AR) and a model for potential evapotranspiration (PET) extracted from the Consortium of
538 Spatial Information (CGIAR-CSI) website (www.cgiar-csi.org). In addition, seven soil
539 variables were extracted from the SOILGRIDS project (<https://soilgrids.org/>, licensed by
540 ISRIC – World Soil Information), downloaded at 250 m resolution and then resampled using
541 the 30 arc second grid of CHELSA (Supplementary Table 2). We refer to these climate and
542 soil data as “environmental data”.

543 **Community trait composition.**

544 For every trait j and plot k , we calculated the plot-level trait means as community-weighted
545 mean (CWM) according to^{2,44}:

$$CWM_{j,k} = \sum_i^{n_k} p_{i,k} t_{i,j}$$

546 where n_k is the number of species sampled in plot k , $p_{i,k}$ is the relative abundance of species i
 547 in plot k , referring to the sum of abundances for all species with traits in the plot, and $t_{i,j}$ is the
 548 mean value of species i for trait j . This computation was done for each of the 18 traits for
 549 1,115,785 plots. The within-plot trait variance is given by community-weighted variance
 550 (CWV)^{44,45}.

$$CWV_{j,k} = \sum_i^{n_k} p_{i,k} (t_{i,j} - CWM_{j,k})^2$$

551 CWV is equal to functional dispersion as described by Rao's quadratic entropy⁴⁶, when using
 552 a squared Euclidean distance matrix $d_{i,j,k}$ ⁴⁷:

$$CWV_{j,k} = \sum_i^{n_k} p_{i,k} (t_{i,j} - CWM_{j,k})^2 = FD_Q = \sum_{i=1}^{n_k-1} \sum_{j=i+1}^{n_k} p_{i,k} p_{j,k} d_{i,j,k}^2$$

553 We had CWV information for 18 traits for 1,099,463 plots, as at least two taxa were needed to
 554 calculate CWV. We performed the calculations using the 'data.table' package⁴⁸ in R.

555

556 **Assessing the degree of filtering.** To analyse how plot-level trait means and within-plot trait
 557 variances (based on gap-filled trait data) depart from random expectation, for each trait we
 558 calculated standardized effect sizes (SESS) for the variance in CWMs and for the mean in
 559 CWVs. Significantly positive SESSs in variance of CWM and significantly negative ones in
 560 the mean of CWV can be considered a global-level measure of environmental or biotic
 561 filtering. To provide an indication of the global direction of filtering, we also report SESSs for
 562 the mean of CWM trait values. Similarly, to measure how much within-community trait
 563 dispersion varied globally, we also calculated SESSs for the variance in CWV.

564 SESSs were obtained from 100 runs of randomizing trait values across all species globally. In
 565 every run we calculated CWM and CWV with random trait values, but keeping all species
 566 abundances in plots. Thus, the results of randomization are independent from species co-
 567 occurrences structure of plots⁴⁹. For every trait, the SESSs of the variance in CWM were
 568 calculated as the observed value of variance in CWM minus the mean variance in CWM of
 569 the random runs, divided by the standard deviation of the variance in CWM of the random
 570 runs (Fig. 1). SESSs for the mean in CWM, the mean in CWV and the variance in CWV were
 571 calculated accordingly. Tests for significance of SESSs were obtained by fitting generalized
 572 Pareto-distribution of the most extreme random values and then estimating p values from this
 573 fitted distribution⁵⁰.

574

575 **Vegetation trait-environment relationships.** Out of the 1,115,785 plots with CWM values,
 576 1,114,304 (99.9%) had complete environmental information and coordinates. This set of plots

577 was used to calculate single linear regressions of each of the 18 traits on each of the 30
578 environmental variables. We used the 'corrplot' function⁵¹ in R to illustrate Pearson
579 correlation coefficients (see Supplementary Figs. 1-2, 4, 6, 8) and for the strongest
580 relationships produced bivariate graphs and mapped the global distribution of the CWM
581 values using kriging interpolation in ArcGIS 10.2 (Fig. 4). We also tested for non-linear
582 relationships with environment by including an additional quadratic term in the linear model
583 and then report coefficients of determination. As in the linear relationships of CWM with
584 environment, the highest r^2 values in models with an additional quadratic term were
585 encountered between stem specific density and PET ($r^2=0.156$) and leaf N:P ratio and
586 growing degree days above 5°C (GDD5, $r^2=0.118$). These were not substantially different
587 from the linear CWM-environment relationships, which had $r^2=0.156$ and $r^2=0.115$,
588 respectively (Fig. 4, Supplementary Fig. 2). Similarly, including a quadratic term in the
589 regressions did not increase the CWV-environment correlations. Here, the strongest
590 correlations were encountered between plant height and soil pH ($r^2=0.044$) and between
591 specific leaf area (SLA) and the volumetric content of coarse fragments in the soil
592 (CoarseFrag, $r^2=0.037$), which were similar to those in the linear regressions ($r^2=0.029$ and
593 $r^2=0.036$, respectively, Supplementary Fig. 3).

594 To account for a possible confounding effect of species richness on CWV, which may cause
595 low CWV through competitive exclusion of species, we regressed CWV on species richness
596 and then calculated all Pearson correlation coefficients with the residuals of this relationship
597 against all climatic variables. Here, the highest correlation coefficients were encountered
598 between PET and CWV of conduit element length ($r^2=0.038$), followed by the relationship of
599 specific leaf area (SLA) and the volumetric content of coarse fragments in the soil
600 (CoarseFrag, $r^2=0.034$), which were very similar in magnitude to the CWV environment
601 correlations ($r^2=0.035$ and $r^2=0.036$, respectively; Supplementary Fig. 3).

602 The CWMs and CWVs were scaled to a mean of zero and standard deviation of one and then
603 subjected to a Principal Component Analysis (PCA), calculated with the 'rda' function from
604 the 'vegan' package⁵². Climate and soil variables were fitted *post hoc* to the ordination scores
605 of plots of the first two axes, producing correlation vectors using the 'envfit' function. We
606 refrain from presenting any inference statistics, as with > 1.1 million plots all environmental
607 variables showed statistically significant correlations. Instead, we report coefficients of
608 determination (r^2), obtained from Redundancy Analysis (RDA), using all 30 environmental
609 variables as constraining matrix, resulting in a maximum of 18 constrained axes
610 corresponding to the 18 traits. We report both r^2 values of the first two axes explained by
611 environment, which is the maximum correlation of the best linear combination of
612 environmental variables to explain the CWM or CWV plot \times trait matrix and r^2 values of all
613 18 constrained axes explained by environment. We plotted the PCA results using the 'ordiplot'
614 function and coloured the points according to the logarithm of the number of plots that fell
615 into grid cells of 0.002 in PCA units (resulting in approximately 100,000 cells). For further
616 details, see the captions of the figures.

617 Additionally, we carried out the PCA and RDA analyses, using CWMs based on original trait
618 values (see above). Because of a poor coverage of some traits we confined the analyses with
619 original trait values to the six traits used by Díaz et al.⁵, which were leaf area, specific leaf

620 area, leaf N, seed mass, plant height and stem specific density. Using these six traits resulted
621 in 954,459 plots that had at least one species with a trait value for each of the six traits.

622

623 **Testing for formation-specific patterns.** We carried out separate analyses for two
624 ‘formations’: forest and for non-forest plots. We defined as forest plots that had > 25% cover
625 of the tree layer. However, this information was available for only 25% of the plots in our
626 sPlot database. Thus, we also assigned formation status based on growth form data from the
627 TRY database. We defined plots as ‘forest’ if the sum of relative cover of all tree taxa was >
628 25%, but only if this did not contradict the requirement of > 25% cover of the tree layer (for
629 those records for which this information was given in the header file). Similarly, we defined
630 non-forest plots by calculating the cover of all taxa that were not defined as trees and shrubs
631 (also taken from the TRY plant growth form information) and that were not taller than 2 m,
632 using the TRY data on mean plant height. We assigned the status ‘non-forest’ to all plots that
633 had >90% cover of these low-stature, non-tree and non-shrub taxa. In total, 21,888 taxa out of
634 the 52,032 in TRY which also occurred in sPlot belonged to this category, and 16,244 were
635 classed as trees. The forests and non-forest plots comprised 330,873 (29.7%) and 513,035
636 (46.0%) of all plots, respectively. We subjected all CWM values for forest and non-forest
637 plots to PCA, RDA and bivariate linear regressions to environmental variables as described
638 above.

639 The forest plots, in particular, confirmed the overall patterns, with respect to variation in
640 CWM explained by the first two PCA axes (60.5%) and the two orthogonal continua from
641 small to large size and the leaf economics spectrum (Supplementary Fig. 6). The variation
642 explained by macroclimate and soil conditions was much larger for the forest subset than for
643 the total data, with the best relationship (leaf N:P ratio and the mean temperature of the
644 coldest quarter, bio11) having $r^2=0.369$ and the second next best ones (leaf N:P ratio and
645 GDD1 and GDD5) close to this value with $r^2=0.357$ (Supplementary Fig. 7) and an overall
646 variation in CWM values explained by environment of 25.3% (cumulative variance of all 18
647 constrained axes in a RDA). The non-forest plots showed the same functional continua, but
648 with lower total amount of variation in CWM accounted for by the first two PCA axes
649 (41.8%, Supplementary Fig. 8) and much lower overall variation explained by environment.
650 For non-forests, the best correlation of any CWM trait with environment was the one of
651 volumetric content of coarse fragments in the soil (CoarseFrag) and leaf C content per dry
652 mass with $r^2=0.042$ (Supplementary Fig. 9). Similarly, the cumulative variance of all 18
653 constrained axes according to RDA was only 4.6%. This shows, on the one hand, that forest
654 and non-forest vegetation are characterized by the same interrelationships of CWM traits, and
655 on the other hand, that the relationships of CWM values with the environment were much
656 stronger for forests than for non-forest formations. The coefficients of determination were
657 even higher than those previously reported for trait-environment relationships for North
658 American forests (between CWM of seed mass and maximum temperature, $r^2=0.281$)³.

659 **Resampling procedure in environmental space.** In order to achieve a more even
660 representation of plots across the global climate space, we first subjected the same 30 global
661 climate and soil variables as described above, to a Principal Component Analysis (PCA),

662 using the climate space of the whole globe, irrespective of the presence of plots in this space,
663 and scaling each variable to a mean of zero and a standard deviation of one. We used a 2.5 arc
664 minute spatial grid, which comprised 8,384,404 terrestrial grid cells. We then counted the
665 number of vegetation plots in the sPlot database that fell into each grid cell. For this analysis,
666 we did not use the full set of 1,117,369 plots with trait information (see above), but only those
667 plots that had a location inaccuracy of max. 3 km, resulting in a total of 799,400 plots. The
668 resulting PCA scores based on the first two principal components (PC1-PC2) were rasterized
669 to a 100 × 100 grid in PC1-PC2 environmental space, which was the most appropriate
670 resolution according to a sensitivity analysis. This sensitivity analysis tested different grid
671 resolutions, from a coarse-resolution bivariate space of 100 grid cells (10 × 10) to a very fine-
672 resolution space of 250,000 grid cells (500 × 500), iteratively increasing the number of cells
673 along each principal component by 10 cells. For each iteration, we computed the total number
674 of sPlot plots per environmental grid cell and plotted the median sampling effort (number of
675 plots) across all grid cells versus the resolution of the PC1-PC2 space. We found that the
676 curve flattens off at a bivariate environmental space of 100 × 100 grid cells, which was the
677 resolution for which the median sampling effort stabilized at around 50 plots per grid cell. As
678 a result, we resampled plots only in environmental cells with more than 50 plots (858 cells in
679 total).

680 To optimize our resampling procedure within each grid cell, we used the heterogeneity-
681 constrained random (HCR) resampling approach⁵³. The HCR approach selects the subset of
682 vegetation plots for which those plots are the most dissimilar in their species composition
683 while avoiding selection of plots representing peculiar and rare communities that differ
684 markedly from the main set of plant communities (outliers), thus providing a representative
685 subset of plots from the resampled grid cell. We used the turnover component of the Jaccard's
686 dissimilarity index (β_{jtu} ⁵⁴) as a measure of dissimilarity. The β_{jtu} index accounts for species
687 replacement without being influenced by differences in species richness. Thus, it reduces the
688 effects of any imbalances that may exist between different plots due to species richness. We
689 applied the HCR approach within a given grid cell by running 1,000 iterations of randomly
690 selecting 50 plots out of the total number of plots available within that grid cell. Where the
691 cell contained 50 or fewer plots, all were included and the resampling procedure was not run.
692 This procedure thinned out over-sampled climate types, while retaining the full environmental
693 gradient.

694 All 1,000 random draws of a given grid cell were subsequently sorted according to the
695 decreasing mean of β_{jtu} between pairs of vegetation plots and then sorted again according to
696 the increasing variance in β_{jtu} between pairs of vegetation plots. Ranks from both sortings
697 were summed for each random draw, and the random draw with the lowest summed rank was
698 considered as the most representative of the focal grid cell. Because of the randomized nature
699 of the HCR approach, this resampling procedure was repeated 100 times for each of the 858
700 grid cells. This enabled us to produce 100 different subsamples out of the full sample of
701 799,400 vegetation plots subjected to the resampling procedure. Each of these 100
702 subsamples was finally subjected to ordinary linear regression, PCA and RDA as described
703 above. We calculated the mean correlation coefficient across the 100 resampled data sets for
704 each environmental variable with each trait.

705 To plot bivariate relationships, we used the mean intercept and slope of these relationships.
706 PCA loadings of all 100 runs were stored and averaged. As different runs showed different
707 orientation on the first PCA axes, we switched the signs of the axis loadings in some of the
708 runs to make the 100 PCAs comparable to the reference PCA, based on the total data set.
709 Across the 100 resampled data sets, we then calculated the minimum and maximum loading
710 for each of the two PCA axes and plotted the result as ellipsoid. We also collected the post-
711 hoc regressions coefficients of PCA scores with the environmental variables in each of the
712 100 runs, switched the signs accordingly and plotted the correlations to PC1 and PC2 as
713 ellipsoids. The result is a synthetic PCA of all 100 runs. To illustrate the coverage of plots in
714 PCA space, we used plot scores of one of the 100 random runs. Similarly, the coefficients of
715 determination obtained from the RDAs of these 100 resampled sets were averaged.

716 The mean PCA loadings across these 100 subsets (summarized in Supplementary Fig. 10)
717 were fully consistent with those of the full data set in Fig. 2, with the same two functional
718 continua in plant size and diaspore mass (from bottom left to top right), and perpendicular to
719 that, the leaf economics spectrum. The variation in CWM accounted for by the first two axes
720 was on average $50.9\% \pm 0.04$ standard deviations (SD), and thus, virtually identical with that
721 in the total dataset. In contrast, the variation explained on average by macroclimate and soil
722 conditions ($26.5\% \pm 0.01$ SD as average cumulative variance of all 18 constrained axes in the
723 RDAs across all 100 runs) was considerably larger than that for the total dataset, which is also
724 reflected in consistently higher correlations between traits and environmental variables
725 (Supplementary Fig. 11). The highest mean correlation was encountered for plant height and
726 PET (mean $r^2=0.342$ across 100 runs). PET was a better predictor for plant height than the
727 precipitation of the wettest months (bio13, mean $r^2=0.231$), as had been suggested
728 previously⁶. The correlation of PET with stem specific density (mean $r^2=0.284$) and warmth
729 of the growing season (expressed as growing degree days above the threshold 5°C , GDD5)
730 with leaf N:P ratio (mean $r^2=0.250$) ranked among the best 12 correlations encountered out of
731 all 540 trait-environment relationships, which confirms the patterns found in the whole data
732 set (compared with Fig. 4). Overall, the coefficients of determination were much closer to the
733 ones reported from other studies with a global collection of a few hundred plots (r^2 values
734 ranging from 36% to 53% based on multiple regressions of single traits with five to six
735 environmental drivers²²).

736

737 **Data availability statement**

738 The data contained in sPlot (the vegetation-plot data complemented by trait and
739 environmental information) are available by request, through contacting any of the sPlot
740 consortium members for submitting a paper proposal. The proposals should follow the
741 Governance and Data Property Rules of the sPlot Working Group, which are available on the
742 sPlot website (www.idiv.de/sPlot).

743

744 **Acknowledgements**

745 sPlot has been initiated by sDiv, the Synthesis Centre of the German Centre for Integrative
746 Biodiversity Research (iDiv) Halle-Jena-Leipzig, funded by the German Research Foundation
747 (FZT 118) and now is a platform of iDiv. H.B., J.De., O.Pu, U.J., B.J.-A., J.K., D.C., F.M.S.,
748 M.W. and C.W. appreciate direct funding through iDiv. For all further acknowledgements see
749 the Supplementary Information.

750

751 **Author contributions**

752 H.B. and U.J. wrote the first draft of the manuscript, with considerable input by B.J.-A. and
753 R.F.; H.B. carried out most of the statistical analyses and produced the graphs; H.B., O.Pu.
754 and U.J. initiated sPlot as an sDiv working group and iDiv platform; J.De. compiled the plot
755 databases globally; J.De., S.M.H., U.J., O.Pu. and F.J. harmonized vegetation databases; J.De.
756 and B.J.-A. coordinated the sPlot consortium; J.K. provided the trait data from TRY; F.S.
757 performed the trait data gap filling; O.Pu. produced the taxonomic backbone; B.J.-A., G.S.
758 and E. Welk compiled environmental data and produced the global maps; S.M.H. wrote the
759 Turboveg v3 software, which holds the sPlot database; J.L. and T.H. wrote the resampling
760 algorithm. Many authors participated in one or more of the three sPlot workshops at iDiv
761 where the sPlot initiative was conceived and planned, and evaluation of the data and first
762 drafts were discussed. All other authors contributed data. All authors contributed to writing
763 the manuscript.

764

765 **Declaration of competing interests**

766 The authors declare no competing interests.

767

768 **References**

- 769 1. Warming, E. *Lehrbuch der ökologischen Pflanzengeographie – Eine Einführung in die*
770 *Kenntnis der Pflanzenvereine*. (Borntraeger, Berlin, 1896).
- 771 2. Garnier, E. *et al.* Plant functional markers capture ecosystem properties during secondary
772 succession. *Ecology* **85**, 2630–2637 (2004).
- 773 3. Ordoñez, J.C. *et al.* A global study of relationships between leaf traits, climate and soil
774 measures of nutrient fertility. *Global Ecol. Biogeogr.* **18**, 137–149 (2009).
- 775 4. Garnier, E., Navas, M.-L. & Grigulis, K. *Plant functional diversity - Organism traits,*
776 *community structure, and ecosystem properties*. (Oxford Univ. Press, 2016).
- 777 5. Díaz, S. *et al.* The global spectrum of plant form and function. *Nature* **529**, 167-171 (2016).
- 778 6. Moles, A.T. *et al.* Global patterns in plant height. *J. Ecol.* **97**, 923-932 (2009).
- 779 7. Wright, I.J. *et al.* The worldwide leaf economics spectrum. *Nature* **428**, 821-827 (2004).

- 780 8. Reich, P.B. The world-wide ‘fast–slow’ plant economics spectrum: a traits manifesto. *J.*
781 *Ecol.* **102**, 275–301 (2014).
- 782 9. Adler, P.B. *et al.* Functional traits explain variation in plant life history strategies. *Proc.*
783 *Natl. Acad. Sci. USA* **111**, 740–745 (2014).
- 784 10. Marks, C.O. & Lechowicz, M.J. Alternative designs and the evolution of functional
785 diversity. *Am. Nat.* **167**, 55–67 (2006).
- 786 11. Grime, J.P. Trait convergence and trait divergence in herbaceous plant communities:
787 Mechanisms and consequences. *J. Veg. Sci.* **17**, 255–260 (2006).
- 788 12. Muscarella, R. & Uriarte, M. Do community-weighted mean functional traits reflect
789 optimal strategies? *Proc. R. Soc. B.* **283**, 20152434 (2016).
- 790 13. Swenson, N.G. & Weiser, M.D. Plant geography upon the basis of functional traits: An
791 example from eastern North American trees. *Ecology* **91**, 2234–2241 (2010).
- 792 14. Fyllas, N.M. *et al.* Basin-wide variations in foliar properties of Amazonian forest:
793 phylogeny, soils and climate. *Biogeosciences* **6**, 2677–2708 (2009).
- 794 15. Swenson, N.G. *et al.* Phylogeny and the prediction of tree functional diversity across
795 novel continental settings. *Global Ecol. Biogeogr.* **26**, 553–562 (2017).
- 796 16. Swenson, N.G. *et al.* The biogeography and filtering of woody plant functional diversity
797 in North and South America. *Global Ecol. Biogeogr.* **21**, 798–808 (2012).
- 798 17. Wright, I.J. *et al.* Global climatic drivers of leaf size. *Science* **357**: 917–921 (2017).
- 799 18. Mayfield, M.M. & Levine, J.M. Opposing effects of competitive exclusion on the
800 phylogenetic structure of communities. *Ecol. Lett.* **13**, 1085 – 1093 (2010).
- 801 19. Kraft, N.J.B. *et al.* Community assembly, coexistence and the environmental filtering
802 metaphor. *Funct. Ecol.* **29**, 592–599 (2015).
- 803 20. Barboni, D. *et al.* Relationships between plant traits and climate in the Mediterranean
804 region: A pollen data analysis. *J. Veg. Sci.* **15**, 635–646 (2004).
- 805 21. Borgy, B. *et al.* Plant community structure and nitrogen inputs modulate the climate signal
806 on leaf traits. *Global Ecol. Biogeogr.* **26**, 1138–1152 (2017).
- 807 22. van Bodegom, P.M, Douma, J.C. & Verheijen, L.M. A fully traits-based approach to
808 modeling global vegetation distribution. *Proc. Natl. Acad. Sci. USA* **111**, 13733–13738
809 (2014).
- 810 23. Moles, A.T. *et al.* Which is a better predictor of plant traits: Temperature or precipitation?
811 *J. Veg. Sci.* **25**, 1167–1180 (2014).
- 812 24. Ordoñez, J.C. *et al.* Plant strategies in relation to resource supply in mesic to wet
813 environments: Does theory mirror nature? *Am. Nat.* **175**, 225–239 (2010).

- 814 25. Simpson, A.J., Richardson, S.J. & Laughlin, D.C. Soil–climate interactions explain
815 variation in foliar, stem, root and reproductive traits across temperate forests. *Global Ecol.*
816 *Biogeogr.* **25**, 964–978 (2016).
- 817 26. Lienin, P. & Kleyer, M. Plant leaf economics and reproductive investment are responsive
818 to gradients of land use intensity. *Agric. Ecosyst. Environ.* **145**, 67–76 (2011).
- 819 27. Maire, V. *et al.* Habitat filtering and niche differentiation jointly explain species relative
820 abundance within grassland communities along fertility and disturbance gradients. *New*
821 *Phytol.* **196**, 497–509 (2012).
- 822 28. Craine, J.M. *et al.* Global patterns of foliar nitrogen isotopes and their relationships with
823 climate, mycorrhizal fungi, foliar nutrient concentrations, and nitrogen availability *New*
824 *Phytol.* **183**, 980–992 (2009).
- 825 29. Güsewell, S. N:P ratios in terrestrial plants: variation and functional significance. *New*
826 *Phytol.* **164**, 243–266 (2004).
- 827 30. Reich, P.B. & Oleksyn, J. Global patterns of plant leaf N and P in relation to temperature
828 and latitude, *Proc. Natl. Acad. Sci. USA* **101**, 11001–11006 (2004).
- 829 31. Scheiter, S., Langan, L. & Higgins, S.I. Next generation dynamic global vegetation
830 models: learning from community ecology. *New Phytol.* **198**, 957–969 (2013).
- 831 32. Boyle, B. *et al.* The Taxonomic Name Resolution Service: an online tool for automated
832 standardization of plant names. *BMC Bioinformatics* **14**, 16 (2013).
- 833 33. Bremer, B. *et al.* An update of the Angiosperm Phylogeny Group classification for the
834 orders and families of flowering plants: APG III. *Bot. J. Linn. Soc.* **161**, 105–121 (2009).
- 835 34. Schrodte, F. *et al.* BHPMF – a hierarchical Bayesian approach to gap-filling and trait
836 prediction for macroecology and functional biogeography. *Global Ecol. Biogeogr.* **24**, 1510–
837 1521 (2015).
- 838 35. Kattge J. *et al.* TRY—a global database of plant traits. *Glob. Change Biol.* **17**, 2905–2935
839 (2011).
- 840 36. Shan, H. *et al.* Gap filling in the plant kingdom - Trait prediction using hierarchical
841 probabilistic matrix factorization. *Proceedings of the 29th International Conference for*
842 *Machine Learning (ICML 2012)* 1303–1310 (2012).
- 843 37. Fazayeli, F. *et al.* Uncertainty quantified matrix completion using Bayesian Hierarchical
844 Matrix factorization. *13th International Conference on Machine Learning and Applications*
845 *(ICMLA 2014)* 312–317 (2014).
- 846 38. Borgy, B. *et al.* Sensitivity of community-level trait–environment relationships to data
847 representativeness: A test for functional biogeography. *Global Ecol. Biogeogr.* **26**, 729–739
848 (2017).

- 849 39. Herz, K. *et al.* Drivers of intraspecific trait variation of grass and forb species in German
850 meadows and pastures. *J. Veg. Sci.* **28**, 705–716 (2017).
- 851 40. Karger, D.N. *et al.* Climatologies at high resolution for the earth’s land surface areas. *Sci.*
852 *Data* **4**, 170122. doi: 10.1038/sdata.2017.122 (2017)
- 853 41. Karger, D.N. *et al.* Climatologies at high resolution for the Earth land surface areas
854 (Version 1.1). *World Data Center for Climate (WDCC) at DKRZ*. [http://chelsa-](http://chelsa-climate.org/downloads/)
855 [climate.org/downloads/](http://chelsa-climate.org/downloads/) (2016).
- 856 42. Dee, D. P. *et al.* The ERA-Interim reanalysis: configuration and performance of the data
857 assimilation system. *Q.J.R. Meteorol. Soc.* **137**, 553–597 (2011).
- 858 43. Synes, N.W. & Osborne, P.E. Choice of predictor variables as a source of uncertainty in
859 continental-scale species distribution modelling under climate change. *Global Ecol. Biogeogr.*
860 **20**, 904–914 (2011).
- 861 44. Enquist, B. *et al.* Scaling from traits to ecosystems: developing a general trait driver
862 theory via integrating trait-based and metabolic scaling theories. *Adv. Ecol. Res.* **52**, 249–318
863 (2015).
- 864 45. Buzzard, V. *et al.* Re-growing a tropical dry forest: functional plant trait composition and
865 community assembly during succession. *Funct. Ecol.* **30**, 1006–1013 (2016).
- 866 46. Rao, C.R. Diversity and dissimilarity coefficients: A unified approach. *Theor. Popul. Biol.*
867 **21**, 24–43 (1982).
- 868 47. Champely, S. & Chessel, D. Measuring biological diversity using Euclidean metrics.
869 *Environ. Ecol. Stat.* **9**, 167–177 (2002).
- 870 48. Dowle, M. *et al.* data.table: Extension of Data.frame. R package version 1.9.6. (2015)
871 <https://CRAN.R-project.org/package=data.table>
- 872 49. Hawkins, B.A. *et al.* Structural bias in aggregated species-level variables driven by
873 repeated species co-occurrences: a pervasive problem in community and assemblage data. *J.*
874 *Biogeogr.* **44**, 1199–1211 (2017).
- 875 50. Knijnenburg, T.A. *et al.* Fewer permutations, more accurate P-values. *Bioinformatics* **25**,
876 i161–i168 (2009).
- 877 51. Friendly, M. Corrgrams: Exploratory displays for correlation matrices. *Am. Statistician*,
878 **56**, 316–324 (2002).
- 879 52. Oksanen, J. *et al.* vegan: Community Ecology Package. R package version 2.3-3 (2016).
880 <https://CRAN.R-project.org/package=vegan>
- 881 53. Lengyel, A., Chytrý, M. & Tichý, L. Heterogeneity-constrained random resampling of
882 phytosociological databases. *J. Veg. Sci.* **22**, 175–183 (2011).

- 883 54. Baselga, A. The relationship between species replacement, dissimilarity derived from
884 nestedness, and nestedness. *Global Ecol. Biogeogr.* **21**, 1223-1232 (2012).
- 885 55. Garnier, E. *et al.* Towards a thesaurus of plant characteristics: an ecological contribution.
886 *J. Ecol.* **105**, 298-309 (2017). www.top-thesaurus.org

887

888

889 **Tables**

890

891 Table 1: Traits used in this study and their function in the community. Traits are arranged
 892 according to the degree to which they should respond to macroclimatic drivers. $\uparrow\downarrow$ in the trait
 893 column denotes opposing relationships, \updownarrow in the description column denotes trade-offs. For
 894 trait units, plot-level trait means and within-plot trait variance see Table 2.

Trait	Description	Function	Expected correlation with macroclimate
Specific leaf area, Leaf area, Leaf fresh mass, Leaf N, Leaf P $\uparrow\downarrow$	Leaf economics spectrum ^{7-8,17} : Thin, N-rich leaves with high turnover and high mass-based assimilation rates \updownarrow	Productivity, competitive ability	Very high ^{12-13,17,21,23}
Leaf dry matter content, Leaf N per area, Leaf C	Thick, N-conservative, long-lived leaves with low mass-based assimilation rates		
Stem specific density	Fast growth \updownarrow Mechanical support, Longevity	Productivity, drought tolerance	Very high ^{12,22}
Conduit element length $\uparrow\downarrow$	Efficient water transport \updownarrow	Water use efficiency	High
Stem conduit density	Safe water transport		
Plant height	Mean individual height of adult plants	Competitive ability	High ^{6,12}
Seed number per reproductive unit $\uparrow\downarrow$	Seed economics spectrum ²³ : Small, well dispersed seeds \updownarrow	Dispersal, regeneration	Moderate ²³⁻²⁴
Seed mass, Seed length, Dispersal unit length	Seeds with storage reserve to facilitate establishment and increase survival		
Leaf N:P ratio	P limitation (N:P > 15) N limitation (N:P < 10) ²⁹	Nutrient supply	Moderate ³⁰
Leaf nitrogen isotope ratio (leaf $\delta^{15}\text{N}$)	Access to N derived from N_2 fixation \updownarrow N supply via mycorrhiza	Nitrogen source, soil depth	Moderate ²⁸

895

896

897 Table 2: Traits, abbreviation of trait names, identifier in the Thesaurus Of Plant characteristics (TOP)⁵⁵, units of measurement, observed values
898 (obs.) standardized effect sizes (SES) and significance (p) of SES for means and variances of both plot-level trait means (community-weighted
899 means, CWMs) and within-plot trait variances (community-weighted variances, CWVs). CWMs and CWVs were based on gap-filled traits for
900 1,115,785 and 1,099,463 plots, respectively. All trait values were log_e-transformed prior to analysis and observed values are on the log_e scale. SES
901 are also based on log_e-transformed values. Stem specific density is stem dry mass per stem fresh volume, specific leaf area is leaf area per leaf dry
902 mass, leaf C, N and P are leaf carbon, nitrogen and phosphorus content, respectively, per leaf dry mass, leaf dry matter content is leaf dry mass per
903 leaf fresh mass, leaf delta ¹⁵N is the leaf nitrogen isotope ratio, stem conduit density is the number of vessels and tracheids per unit area in a cross
904 section, conduit element length refers to both vessels and tracheids. SESs were calculated by randomizing trait values across all species globally 100
905 times and calculating CWM and CWV with random trait values, but keeping all species abundances in plots (see Fig. 1). Tests for significance of
906 SES were obtained by fitting generalized Pareto-distribution of the most extreme random values and then estimating p values from this fitted
907 distribution⁵⁰. * indicates significance at *p* < 0.05.

Trait	Abbreviation	TOP	Unit	CWM						CWV					
				mean			variance			mean			variance		
				obs.	SES	<i>p</i>	obs.	SES	<i>p</i>	obs.	SES	<i>p</i>	obs.	SES	<i>p</i>
Leaf area	LA	25	mm ²	6.130	-9.75	*	1.691	12.53	*	1.565	-2.59	*	2.448	-0.27	n.s.
Specific leaf area	SLA	50	m ² kg ⁻¹	2.850	9.89	*	0.172	12.88	*	0.150	-1.33	n.s.	0.023	1.10	n.s.
Leaf fresh mass	Leaf.fresh.mass	35	g	-2.125	-13.28	*	1.395	10.83	*	1.520	-2.05	*	2.311	0.01	n.s.
Leaf dry matter content	LDMC	45	g g ⁻¹	-1.294	-5.67	*	0.101	11.52	*	0.130	0.95	n.s.	0.017	6.73	*
Leaf C	LeafC	452	mg g ⁻¹	6.116	-3.77	*	0.003	8.80	*	0.002	-1.78	*	0.000	-0.38	n.s.
Leaf N	LeafN	462	mg g ⁻¹	3.038	4.22	*	0.055	6.29	*	0.063	-3.19	*	0.004	-0.13	n.s.
Leaf P	LeafP	463	mg g ⁻¹	0.535	9.57	*	0.097	2.81	*	0.117	-5.17	*	0.014	-2.11	*
Leaf N per area	LeafN.per.area	481	g m ⁻²	0.251	-9.06	*	0.075	8.18	*	0.099	-0.28	n.s.	0.010	1.54	n.s.
Leaf N:P ratio	Leaf.N:P.ratio	-	g g ⁻¹	2.444	-11.95	*	0.040	0.40	n.s.	0.081	-2.74	*	0.007	-0.39	n.s.
Leaf δ ¹⁵ N	Leaf.delta15N	-	ppm	0.521	-3.58	*	0.254	6.68	*	0.455	2.82	*	0.207	2.44	*
Seed mass	Seed.mass	103	mg	0.407	-11.19	*	2.987	3.69	*	2.784	-9.06	*	7.750	-2.81	*
Seed length	Seed.length	91	mm	1.069	-4.51	*	0.294	5.50	*	0.365	-4.67	*	0.134	-3.07	*
Seed number per reproductive unit	Seed.num.rep.unit	-		6.179	7.67	*	2.783	4.40	*	5.156	1.44	n.s.	26.588	2.25	*
Dispersal unit length	Disp.unit.length	90	mm	1.225	-2.51	*	0.343	6.50	*	0.451	-3.21	*	0.203	-1.39	n.s.

Plant height	Plant.height	68	m	-0.315	-12.15	*	1.532	13.34	*	1.259	-9.01	*	1.585	9.68	*
Stem specific density	SSD	286	g cm ⁻³	-0.869	-14.93	*	0.041	13.15	*	0.058	2.09	*	0.003	2.99	*
Stem conduit density	Stem.cond.dens	-	mm ⁻²	4.407	15.08	*	0.656	8.45	*	0.975	-0.95	n.s.	0.951	1.10	n.s.
Conduit element length	Cond.elem.length	-	μm	5.946	-7.09	*	0.182	9.14	*	0.367	7.12	*	0.135	5.29	*
Mean SES					-3.50			8.06			-1.76			1.25	
Mean absolute SES					8.66			8.06			3.36			2.43	

908

909

910 **Captions of Figures**

911

912 Fig. 1: Conceptual figure to illustrate Hypothesis 1, stating that environmental or biotic
913 filtering of community trait values result in a) higher than expected variation of community-
914 weighted means and b) lower than expected community-weighted variances of trait values.
915 Both figures give an example for a single trait and show the relative abundance of trait values
916 of all species in a plot. Black curves refer to observed plot-level trait values in two exemplary
917 plots, while grey curves show plot-level trait values obtained from randomizing trait values
918 across all species globally (see Methods). Randomization was done 100 times, but only one
919 randomization event is shown. Deviation from random expectation was assessed with
920 standardized effect sizes (SEs) for a) the variance in CWMs and b) for the mean in CWVs.
921 Evidence for filtering is given in a) if the variance in plot-level trait means was higher than
922 expected by chance (SES significantly positive) or b) if within-plot trait variance was
923 typically lower than expected by chance (SES significantly negative, see Methods).

924 Fig. 2: Principal Component Analysis of global plot-level trait means (community-weighted
925 means, CWMs). The plots ($n=1,114,304$) are shown by coloured dots, with shading indicating
926 plot density on a logarithmic scale, ranging from yellow with 1–4 plots at the same position to
927 dark red with 251–1142 plots. Prominent spikes are caused by a strong representation of
928 communities with extreme trait values, such as heathlands with ericoid species with small leaf
929 area and seed mass. Post-hoc correlations of PCA axes with climate and soil variables are
930 shown in blue and magenta, respectively. Arrows are enlarged in scale to fit the size of the
931 graph; thus, their lengths show only differences in variance explained relative to each other.
932 Variance in CWM explained by the first and second axis was 29.7% and 20.1%, respectively.
933 The vegetation sketches schematically illustrate the size continuum (short *vs.* tall) and the leaf
934 economics continuum (low *vs.* high LDMC and leaf N content per area in light and dark green
935 colours, respectively). See Table 2 and Supplementary Table 2 for the description of traits and
936 environmental variables.

937

938 Fig. 3: Principal Component Analysis of global within-plot trait variances (community-
939 weighted variances, CWVs). The plots ($n=1,098,015$) are shown by coloured dots, with
940 shading indicating plot density on a logarithmic scale, ranging from yellow with 1–2 plots at
941 the same position to dark red with 631–1281 plots. Post-hoc correlations of PCA axes with
942 climate and soil variables are shown in blue and magenta, respectively. Arrows are enlarged
943 in scale to fit the size of the graph; thus, their lengths show only differences in variance
944 explained relative to each other. Variance in CWV explained by the first and second axis was
945 24.9% and 13.4%, respectively. CWV values of all traits increased from the left to the right,
946 which reflects increasing species richness ($r^2 = 0.116$ between scores of the first axis and
947 number of species in the communities for which traits were available). The vegetation
948 sketches schematically illustrate low and high variation in the plant size and leaf economics
949 continua. See Table 2 and Supplementary Table 2 for the description of traits and
950 environmental variables.

951

952 Fig. 4: The two strongest relationships found for global plot-level trait means (community-
953 weighted means, CWMs) in the sPlot dataset. CWM of the natural logarithm of stem specific
954 density [g cm^{-3}] as a) global map, interpolated by kriging within a radius of 50 km around the
955 plots using a grid cell of 10 km, and b) function of potential evapotranspiration (PET,
956 $r^2=0.156$). CWM of the natural logarithm of the N:P ratio [g g^{-1}] as c) global kriging map and
957 d) function of the warmth of the growing season, expressed as growing degree days over a
958 threshold of 5°C (GDD5, $r^2=0.115$). Plots with N:P ratios > 15 (of 2.71 on the \log_e scale) tend
959 to indicate phosphorus limitation²⁹ and are shown above the broken line in red colour (90,979
960 plots, 8.16% of all plots). The proportion of plots with N:P ratios > 15 increases with GDD5
961 ($r^2=0.895$ for a linear model on the log response ratio of counts of plots with N:P > 15 and
962 ≤ 15 counted within bins of 500 GDD5).

963





