Global trait–environment relationships of plant communities

Bruelheide, H.; Dengler, J.; Purschke, O.; Lenoir, J.; Jiménez-Alfaro, B.; Hennekens, S.M.; ... ; Jandt, U.

Article details

Bruelheide H., Dengler J., Purschke O., Lenoir J., Jiménez-Alfaro B., Hennekens S.M., Botta-Dukát Z., Chytrý M., Field R., Jansen F., Kattge J., Pillar V.D., Schrodt F., Mahecha M.D., Peet R.K., Bodegom P.M. van, Sandel B., Alvarez-Dávila E., Altman J., Attorre F., Arfin Khan M.A.S., Baraloto C., Aubin I., Bauters M., Barroso J.G., Biurrun I., Bergmeier E., Blonder B., Bjorkman A.D., Cayuela L., Carni A., Craven D., Dainese M., Cerný T., Cornelissen J.H.C., Díaz S., Doležal J., Derroire G., De Sanctis M., Fenton N.J., Garnier E., Farfan-Rios W., Feldpausch T.R., Haider S., Hattab T., Guerin G.R., Gutiérrez A.G., Hölzel N., Higuchi P., Hérault B., Henry G., Kacki Z., Jürgens N., Jentsch A., Homeier J., Knollová I., Kleyer M., Kessler M., Karger D.N., Lens F., Laughlin D.C., Kühn I., Korolyuk A.Y., Marcenò C., Mencuccini M., Müller J.V., Munzinger J., Loos J., Louault F., Lyubenova M.I., Malhi Y., Ozinga W.A., Penuelas J., Pérez-Haase A., Petrík P., Myers-Smith I.H., Neill D.A., Niinemets Ü., Orwin K.H., Sabatini F.M., Rodrigues A.V., Schmidt M., Sardans J., Pärtel M., Phillips OL., Römermann C., Reich B., Svenning J.C., Sporbert M., Thomas R., Tang Z., Silva Espejo J.E., Seidler G., Smyth A., Silveira M., Wesche K., Winter M., Weiher E., Welk E., Violle C., Virtanen R., Tsiripidis I., Vassilev K., Wirth C. & Jandt U.

(2018), Global trait–environment relationships of plant communities, Nature ecology & evolution 2(12): 1906-1917.
Doi: 10.1038/s41559-018-0699-8

nature ecology & evolution

ARTICLES

https://doi.org/10.1038/s41559-018-0699-8

Global trait–environment relationships of plant communities

Helge Bruelheide[1,2*], Jürgen Dengler[2,3,4], Oliver Purschke[1,2], Jonathan Lenoir[5], Borja Jiménez-Alfaro[6,1,2], Stephan M. Hennekens[7], Zoltán Botta-Dukát[8], Milan Chytrý[9], Richard Field[10], Florian Jansen[11], Jens Kattge[2,12], Valério D. Pillar[13], Franziska Schrodt[10,12], Miguel D. Mahecha[2,12], Robert K. Peet[14], Brody Sandel[15], Peter van Bodegom[16], Jan Altman[17], Esteban Alvarez-Dávila[18], Mohammed A. S. Arfin Khan[19,20], Fabio Attorre[21], Isabelle Aubin[22], Christopher Baraloto[23], Jorcely G. Barroso[24], Marijn Bauters[25], Erwin Bergmeier[26], Idoia Biurrun[27], Anne D. Bjorkman[28], Benjamin Blonder[29,30], Andraž Čarni[31,32], Luis Cayuela[33], Tomáš Černý[34], J. Hans C. Cornelissen[35], Dylan Craven[2,36], Matteo Dainese[37], Géraldine Derroire[38], Michele De Sanctis[21], Sandra Díaz[39], Jiří Doležal[17], William Farfan-Rios[40,41], Ted R. Feldpausch[42], Nicole J. Fenton[43], Eric Garnier[44], Greg R. Guerin[45], Alvaro G. Gutiérrez[46], Sylvia Haider[1,2], Tarek Hattab[47], Greg Henry[48], Bruno Hérault[49,50], Pedro Higuchi[51], Norbert Hölzel[52], Jürgen Homeier[53], Anke Jentsch[20], Norbert Jürgens[54], Zygmunt Kącki[55], Dirk N. Karger[56,57], Michael Kessler[56], Michael Kleyer[58], Ilona Knollová[9], Andrey Y. Korolyuk[59], Ingolf Kühn[36,1,2], Daniel C. Laughlin[60,61], Frederic Lens[62], Jacqueline Loos[63], Frédérique Louault[64], Mariyana I. Lyubenova[65], Yadvinder Malhi[66], Corrado Marcenò[27], Maurizio Mencuccini[67,68], Jonas V. Müller[69], Jérôme Munzinger[70], Isla H. Myers-Smith[71], David A. Neill[72], Ülo Niinemets[73], Kate H. Orwin[74], Wim A. Ozinga[7,75], Josep Penuelas[68,73,76], Aaron Pérez-Haase[77,78], Petr Petřík[17], Oliver L. Phillips[79], Meelis Pärtel[80], Peter B. Reich[81,82], Christine Römermann[2,83], Arthur V. Rodrigues[84], Francesco Maria Sabatini[1,2], Jordi Sardans[68,76], Marco Schmidt[85], Gunnar Seidler[1], Javier Eduardo Silva Espejo[86], Marcos Silveira[87], Anita Smyth[45], Maria Sporbert[1,2], Jens-Christian Svenning[28], Zhiyao Tang[88], Raquel Thomas[89], Ioannis Tsiripidis[90], Kiril Vassilev[91], Cyrille Violle[44], Risto Virtanen[2,92,93], Evan Weiher[94], Erik Welk[1,2], Karsten Wesche[2,95,96], Marten Winter[2], Christian Wirth[2,12,97] and Ute Jandt[1,2]

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

How climate drives the functional characteristics of vegetation across the globe has been a key question in ecological research for more than a century[1]. While functional information is available for a large portion of the global pool of plant species, we do not know how functional traits of the different species that co-occur in a community are combined, which is what

A full list of affiliations appears at the end of the paper.

Table 1 | Traits used in this study and their function in the community

Trait	Description	Function	Expected correlation with macroclimate
Specific leaf area, Leaf area, Leaf fresh mass, Leaf N, Leaf P	Leaf economics spectrum[7,8,17]: Thin, N-rich leaves with high turnover and high mass-based assimilation rates	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⇔Mechanical support, Longevity	Productivity, drought tolerance	Very high[12,22]
Conduit element length	Efficient water transport	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	Seed economics spectrum[23]: Small, well-dispersed seeds	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)	Nutrient supply	Moderate[30]
	N limitation (N/P < 10)[29]		
Leaf nitrogen isotope ratio (leaf $\delta^{15}N$)	Access to N derived from N_2 fixation⇔N supply via mycorrhiza	Nitrogen source, soil depth	Moderate[28]

Traits are arranged according to the degree to which they should respond to macroclimatic drivers. ↑↓ in the trait column denotes opposing relationships, ⇳ in the description column denotes trade-offs. For trait units, plot-level trait means and within-plot trait variance see Table 2.

determines their joint effect on ecosystems[2–4]. At the species level, Díaz et al.[5] demonstrated that 74% of the global spectrum of six key plant traits determining plant fitness in terms of survival, growth and reproduction can be accounted for by two principal components. They showed that the functional space occupied by vascular plant species is strongly constrained by trade-offs between traits and converges on a small set of successful trait combinations, confirming previous findings[6–9]. While these constraints describe evolutionarily viable ecological strategies for vascular plant species globally, they provide only limited insight into trait composition within communities. There are many reasons why trait composition within communities would produce very different patterns, and indeed much theory predicts this[10,11]. However, it is still unknown to what extent community-level trait composition depends on local factors (microclimate, fine-scale soil properties, disturbance regime[10], successional dynamics[2]) and regional to global environmental drivers (macroclimate[6,12,13], coarse-scale soil properties[3,14]). As ecosystem functions and services are ultimately dependent on the traits of the species composing ecological communities, exploring community trait composition at the global scale can advance our understanding of how climate change and other anthropogenic drivers affect ecosystem functioning.

So far, studies relating trait composition to the environment at continental to global extents have been restricted to coarse-grained species occurrence data (for example, presence in 1° grid cells[15–17]). Such data capture neither biotic interactions (co-occurrence in large grid cells does not indicate local co-existence), nor local variation in environmental filters (for example, variation in soil, topography or disturbance regime within grid cells). In contrast, functional composition of ecological communities sampled at fine-grained vegetation plots—with areas of a few to a few hundred square metres—is the direct outcome of the interaction between both local and large-scale factors. Here, we present a global analysis of plot-level trait composition. We combined the 'sPlot' database, a new global initiative incorporating more than 1.1 million vegetation plots from over 100 databases (mainly forests and grasslands; see Methods), with 30 large-scale environmental variables and 18 key plant functional traits derived from TRY, a global plant-trait database (see Methods and Table 2). We selected these 18 traits because they affect different key ecosystem processes and are expected to respond to macroclimatic drivers (Table 1). In addition, they were sufficiently measured across all species globally to allow for imputation of missing values (see Methods). All analyses were confined to vascular plant species and included all vegetation layers in a community, from the canopy to the herb layer (see Methods).

We used this unprecedented fine-resolution dataset to test the hypothesis (Hypothesis 1) that plant communities show evidence of environmental or biotic filtering at the global scale, making use of the observed variation of plot-level trait means and means of within-plot trait variation across communities. Ecological theory suggests that community-level convergence could be interpreted as the result of filtering processes, including environmental filtering and biotic interactions. Globally, temperature and precipitation drive the differences in vegetation between biomes, suggesting strong environmental filtering[3,11] that constrains the number of successful trait combinations and leads to community-level trait convergence. Similarly, biotic interactions may eliminate excessively divergent trait combinations[18,19]. However, alternative functional trait combinations may confer equal fitness in the same environment[10]. If plant communities show a global variation of plot-level trait means higher than expected by chance, and a lower than expected within-plot trait variation (see Fig. 1), this would support the view that environmental or biotic filtering are dominant structuring processes of community trait composition at the global scale. A consequence of strong community-level trait convergence, and thus low variation within plots with species trait values centred around the mean, would be that plot-level means will be similar to the trait values of the species in that plot. Hence, community mean trait values should mirror the trait values of individual species[5].

Table 2 | Traits, abbreviation of trait names, identifier in the Thesaurus Of Plant characteristics (TOP; www.top-thesaurus.org, ref. [55]), units of measurement, observed values (obs.) SESs and significance (*P*) of SES for means and variances of both plot-level trait means (CWMs) and within-plot trait variances (CWVs)

				CWM						CWV					
				Mean			Variance			Mean			Variance		
Trait	**Abbreviation**	**TOP**	**Unit**	**Obs.**	**SES**	***P***	**Obs.**	**SES**	***P***	**Obs.**	**SES**	***P***	**Obs.**	**SES**	***P***
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 $\delta^{15}N$	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	

CWMs and CWVs were based on gap-filled traits for 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. SESs 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 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 leaf fresh mass, leaf delta ^{15}N is the leaf nitrogen isotope ratio, stem conduit density is the number of vessels and tracheids per unit area in a cross section, conduit element length refers to both vessels and tracheids. SESs were calculated by randomizing trait values across all species globally 100 times and calculating CWM and CWV with random trait values, but keeping all species abundances in plots (see Fig. 1). Tests for significance of SESs were obtained by fitting generalized Pareto-distribution of the most extreme random values and then estimating *P* values from this fitted distribution[50]. * indicates significance at $P < 0.05$.

While Hypothesis 1 addresses the degree of filtering, it does not make a statement on the attribution of driving factors. The main drivers should correlate strongly (though not necessarily linearly[20]) with plot-level trait means and within-plot trait variance. Identifying these drivers has the potential to fundamentally improve our understanding of global trait–environment relationships. We tested the hypothesis (Hypothesis 2) that there are strong correlations between global environmental drivers such as macroclimate and coarse-scale soil properties and both plot-level trait means and within-plot trait variances[3,6,12–17,20–24] (see Table 1 for expected relationships and Supplementary Table 2 for variables used). Such evidence, although correlative, may contribute to the formulation of novel hypotheses to explain global plant trait patterns.

Results and discussion

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

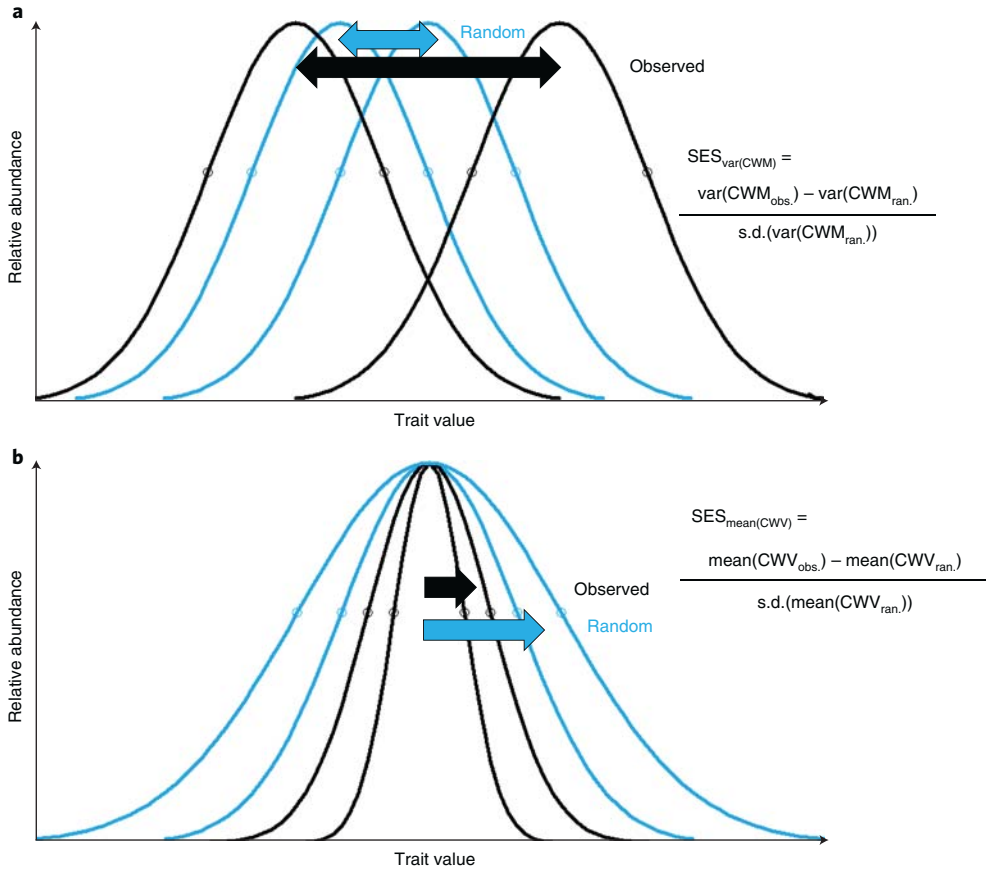

Fig. 1 | Conceptual figure to illustrate Hypothesis 1. a,**b**, Environmental or biotic filtering of community trait values result in higher than expected variation of community-weighted means (**a**) and lower than expected community-weighted variances of trait values (**b**). Both **a** and **b** give an example for a single trait and show the relative abundance of trait values of all species in a plot. Black curves refer to observed plot-level trait values in two exemplary plots, while blue curves show plot-level trait values obtained from randomizing trait values across all species globally (see Methods). Randomization was done 100 times, but only one randomization event is shown. Deviation from random expectation was assessed with SESs for the variance in CWMs (**a**) and for the mean in CWVs (**b**). Evidence for filtering is given in **a** if the variance in plot-level trait means was higher than expected by chance (SES significantly positive) or in **b** if within-plot trait variance was typically lower than expected by chance (SES significantly negative, see Methods).

Trait correlations at the community level were relatively well captured by the first two axes of a principal component analysis (PCA) for both plot-level trait means and within-plot trait variances (Figs. 2 and 3). The dominant axes were determined by those traits with the highest absolute SESs of plot-level trait mean trait values (Table 2, mean of community-weighted means (CWMs)). The PCA of plot-level trait means (Fig. 2) reflects two main functional continua on which community trait values converge: one from short-stature, small-seeded communities such as grasslands or herbaceous vegetation to tall-stature communities with large, heavy diaspores such as forests (the size spectrum), and the other from communities with resource-acquisitive to those with resource-conservative leaves (that is the leaf economics spectrum)[7]. The high similarity between this PCA and the one at the species level by Díaz et al.[5] is striking. Here, at the community level, based on 1.1 million plots, the same functional continua emerged as at the species level, based on 2,214 species. While the trade-offs between different traits at the species level can be understood from a physiological and evolutionary perspective, finding similar trade-offs between traits at the community level was unexpected, as species with opposing trait values can co-exist in the same community. In combination with our finding of strong trait convergence, these results reveal a strong parallel of present-day community assembly with individual species' evolutionary histories.

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

The ordination of within-plot variance of the different traits (Fig. 3) revealed two main continua. Variances of plant height and diaspore mass varied largely independently of variances of traits representing the leaf economics spectrum. This suggests that short and tall species can be assembled together in the same community independently from combining species with acquisitive leaves with species with conservative leaves. Global climate and soil variables accounted for even less variation on the first two PCA axes in within-plot trait variances than on the first two PCA axes

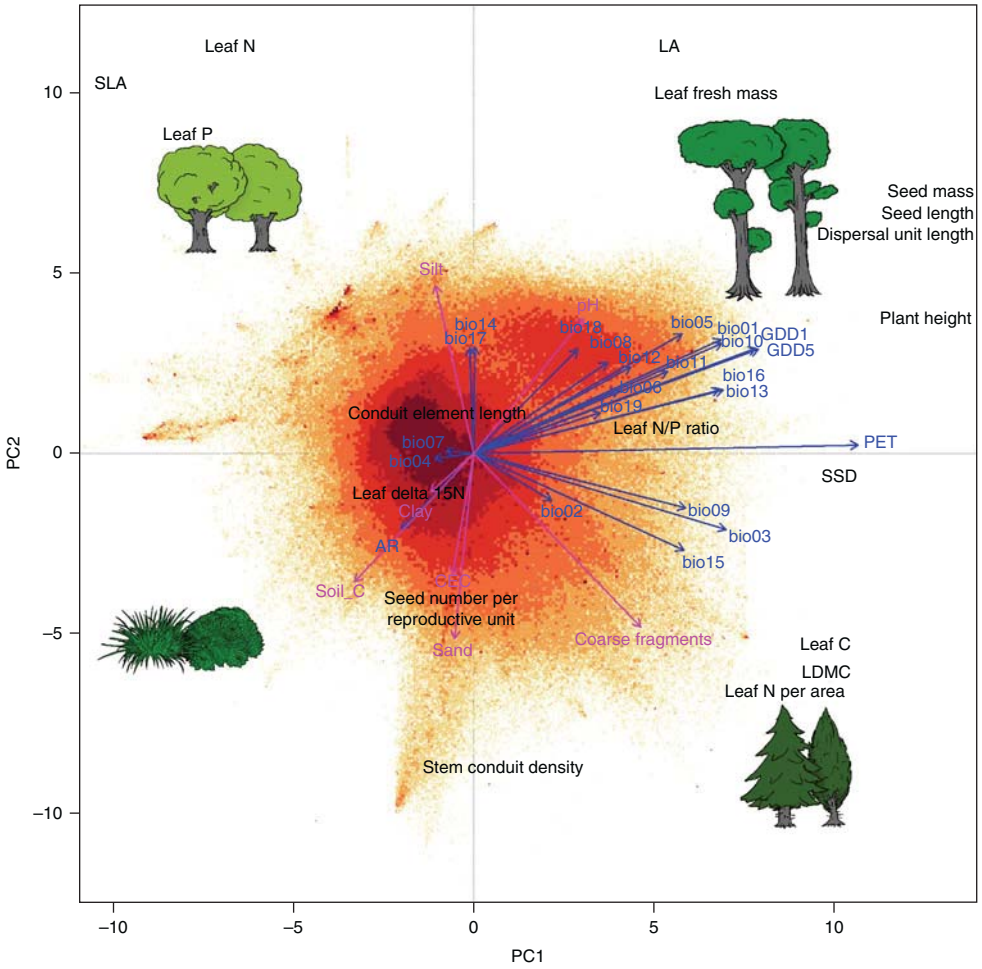

Fig. 2 | Principal component analysis of global plot-level trait means (CWMs). The plots (n = 1,114,304) are shown by coloured dots, with shading indicating plot density on a logarithmic scale, ranging from yellow with 1–4 plots at the same position to dark red with 251–1,142 plots. Prominent spikes are caused by a strong representation of communities with extreme trait values, such as heathlands with ericoid species with small leaf area and seed mass. Post hoc correlations of PCA axes with climate and soil variables are shown in blue and magenta, respectively. Arrows are enlarged in scale to fit the size of the graph; thus, their lengths show only differences in variance explained relative to each other. Variance in CWM explained by the first and second axes was 29.7% and 20.1%, respectively. The vegetation sketches schematically illustrate the size continuum (short versus tall) and the leaf economics continuum (low versus high leaf dry matter content (LDMC) and leaf N content per area in light and dark green colours, respectively). See Table 2 and Supplementary Table 2 for the description of traits and environmental variables.

in plot-level trait means. Only two environmental variables had $r^2 > 3\%$ (Supplementary Fig. 3), whether allowing for non-linear relationships (see Methods) or not, and, overall, macro-environment accounted for only 3.6% or 5.0% of the variation (cumulative variance, respectively, of the first two or all 18 constrained axes). Removing species richness effects from within-plot trait variances did not increase the amount of variation explained by the environment (see Methods).

The findings of our study contrast strongly with studies where the variation in traits between species was calculated at the level of the species pool in large grid cells[15,16], suggesting that plot-level and grid cell-level trait composition are driven by different factors[21]. Plot-level trait means and variances may both be predominantly driven by local environmental factors, such as topography (for example north-facing versus south-facing slopes), local soil characteristics (for example soil depth and nutrient supply)[3,14,24,25], disturbance regime (including land use[26] and successional status[2,27]) and biotic interactions[18,19,28], while broad-scale climate and soil conditions may only become relevant for the whole species pool in large grid cells. Such differences emphasize the importance of the effect of local environment on communities' trait composition, which should be taken into account when interpreting the effect of environmental drivers on functional trait diversity, using data on either floristic pools or ecological communities.

We note that the strongest community-level correlations with environment were found for traits not linked to the leaf economics spectrum. Mean stem specific density increased with potential evapotranspiration (PET, $r^2 = 15.6\%$; Fig. 4a,b), reflecting the need to produce denser wood with increasing evaporative demand. Leaf N/P ratio increased with growing-season warmth (growing degree days above 5 °C, GDD5, $r^2 = 11.5\%$; Fig. 4d), indicating strong phosphorus limitation[29] in most plots in the tropics and subtropics (Fig. 4c,d). This pattern was not brought about by a parallel increase in the presence of legumes, which tend to have relatively high N/P ratios; excluding all species of Fabaceae resulted in a very similar relationship with GDD5 ($r^2 = 10.0\%$). The global N/P pattern is consistent with results based on traits of single species related to mean annual temperature[30]. We assume that the main underlying mechanism is the high soil weathering rate at high temperatures and humidity, which in the tropics and subtropics was not reset by Pleistocene glaciation. Thus, phosphorus limitation may weaken the relationships between productivity-related traits and macroclimate

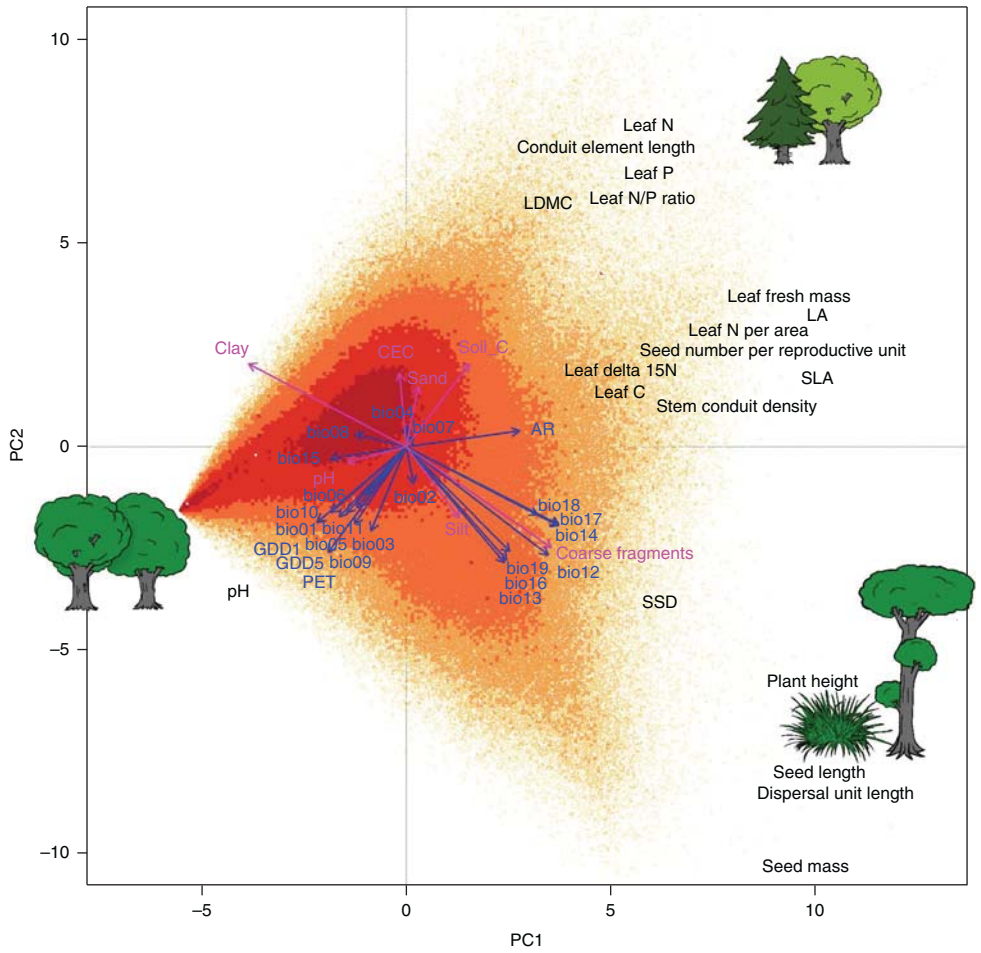

Fig. 3 | Principal component analysis of global within-plot trait variances (CWVs). The plots ($n = 1{,}098{,}015$) are shown by coloured dots, with shading indicating plot density on a logarithmic scale, ranging from yellow with 1–2 plots at the same position to dark red with 631–1,281 plots. Post hoc correlations of PCA axes with climate and soil variables are shown in blue and magenta, respectively. Arrows are enlarged in scale to fit the size of the graph; thus, their lengths show only differences in variance explained relative to each other. Variance in CWV explained by the first and second axes was 24.9% and 13.4%, respectively. CWV values of all traits increase from the left to the right, which reflects increasing species richness ($r^2 = 0.116$ between scores of the first axis and number of species in the communities for which traits were available). The vegetation sketches schematically illustrate low and high variation in the plant size and leaf economics continua. See Table 2 and Supplementary Table 2 for the description of traits and environmental variables.

(Supplementary Fig. 2). For example, specific leaf area (SLA) may be low as a consequence of low nutrient availability[3,14,24,25] in favourable climates as well as being low as a consequence of low temperature and precipitation with favourable nutrient supply. Overall, our findings are relevant for improving dynamic global vegetation models (DGVMs), which so far have used trait information only from a few calibration plots[22]. The sPlot database provides much-needed empirical data on the community trait pool in DGVMs[31] and identifies traits that should be considered when predicting ecosystem functions from vegetation, such as stem specific density and leaf N/P ratio.

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

We also assessed whether the observed trait–environment relationships hold for forest and non-forest vegetation independently (see Methods). Both subsets confirmed the overall patterns in trait means (Supplementary Figs. 3–6). The variance in plot-level trait means explained by large-scale climate and soil variables was higher for forest than non-forest plots, probably because forests belong to a well-defined and rather resource-conservative formation, whereas non-forest plots encompass a heterogeneous mixture of different vegetation types, ranging from alpine meadows to semi-deserts, and tend to depend more on disturbance and management, which can strongly affect trait–environment relationships of communities[21]. Finally, to test whether our findings depended on the uneven distribution of plots among the world's different climates and soils, we

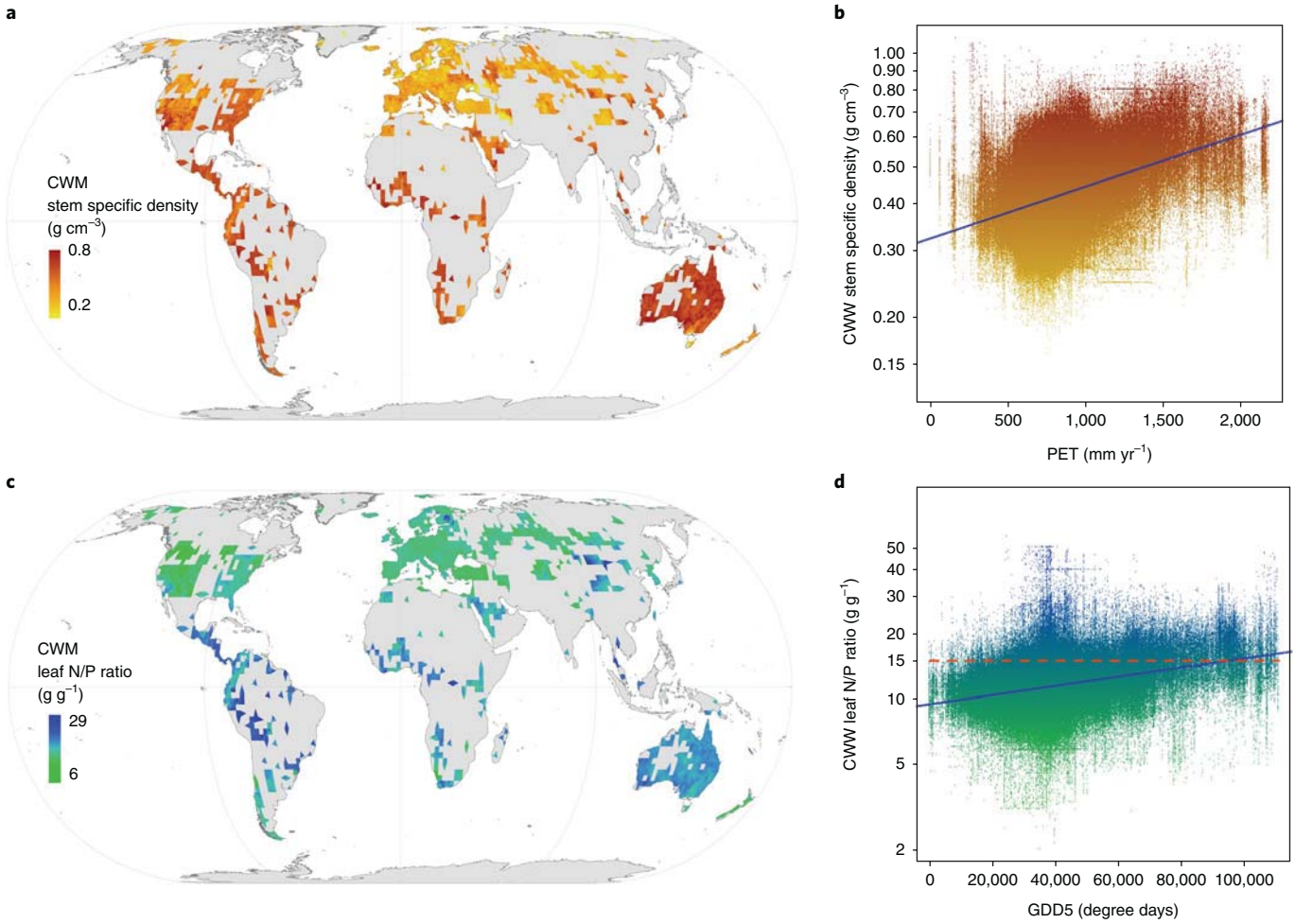

Fig. 4 | The two strongest relationships found for global plot-level trait means (CWMs) in the sPlot dataset. a,**b**, CWM of the natural logarithm of stem specific density (g cm^{-3}) as a global map, interpolated by kriging within a radius of 50 km around the plots using a grid cell of 10 km (**a**), and function of potential evapotranspiration (PET, $r^2 = 0.156$) (**b**). **c**,**d**, CWM of the natural logarithm of the N/P ratio (g g^{-1}) as a global kriging map (**c**) and function of the warmth of the growing season, expressed as growing degree days over a threshold of 5 °C (GDD5, $r^2 = 0.115$) (**d**). Plots with N/P ratios > 15 (of 2.71 on the $\log_e$ scale) tend to indicate phosphorus limitation[29] and are shown above the broken line in red colour (90,979 plots, 8.16% of all plots). The proportion of plots with N/P ratios >15 increases with GDD5 ($r^2 = 0.895$ for a linear model on the log response ratio of counts of plots with N/P > 15 and ≤15 counted within bins of 500 GDD5).

repeated the analyses in 100 subsets of ~100,000 plots resampled in the global climate space (Supplementary Figs. 7 and 8). The analysis of the resampled datasets revealed the same patterns and confirmed the impact of PET and GDD5 on stem specific density and leaf N/P ratio, respectively. The correlations between trait means and environmental variables were, however, stronger in the resampled subsets, possibly because the resampling procedure reduced the overrepresentation of the temperate-zone areas with intermediate climatic values.

Our findings have important implications for understanding and predicting plant community trait assembly. First, worldwide trait variation of plant communities is captured by a few main dimensions of variation, which are surprisingly similar to those reported by species-level studies[5,7–9], suggesting that the drivers of past trait evolution, which resulted in the present-day species-level trait spectra[5], are also reflected in the composition of today's plant communities. If species-level trade-offs indeed constrain community assembly, then the present-day contrasts in trait composition of terrestrial plant communities should also have existed in the past and will probably remain, even for novel communities, in the future. Most species in our present-day communities evolved under very variable filtering conditions across the globe, with respect to temperature and precipitation regimes. Therefore, it can be assumed that future filtering conditions will result in novel communities that follow the same functional continua from short-stature, small-seeded communities to tall-stature communities with large, heavy diaspores and from communities with resource-acquisitive leaves to those with resource-conservative leaves. Second, the main plot-level vegetation trait continua cannot easily be captured by coarse-resolution environmental variables[21]. This brings into question both the use of simple large-scale climate relationships to predict the global spectra of plant assemblages[13,15,16,22] and attempts to derive net primary productivity and global carbon and water budgets from global climate, even when employing powerful trait-based vegetation models[31]. The finding that within-plot trait variance is only very weakly related to global climate or soil variables points to the importance of i) local-scale climate or soil variables, ii) disturbance regimes and iii) biotic interactions for the degree of local trait dispersion[11]. Finally, our findings on the limited role of large-scale climate in explaining trait patterns and on the prevalence of phosphorus limitation in most plots in the tropics and subtropics call for including local variables when predicting community trait

patterns. Even under similar macro-environmental conditions, communities can vary greatly in trait means and variances, consistent with high local variation in species' trait values[3,6,7]. Future research on the functional response of communities to changing climate should incorporate the effect of local environmental conditions[24–26] and biotic interactions[18,19] in building reliable predictions of vegetation dynamics.

Methods

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

For most plots (97.2%), information on the single species' relative contribution to the sum of plants in the plot was available, expressed as cover, basal area, individual count, importance value or percentage frequency in subplots. For the other 2.8% (31,461 plots), for which only presence/absence (*p*/*a*) was available, we assigned equal relative abundance to the species (1/species richness). For plots with a mix of cover and *p*/*a* information (mostly forest plots, where herb layer information had been added on a *p*/*a* basis; 8,524 plots), relative abundance was calculated by assigning the smallest cover value that occurred in a particular plot to all species with only *p*/*a* information in that plot. In most cases (98.4%), plot records in sPlot include full species lists of vascular plants. Bryophytes and lichens were additionally identified in 14% and 7% of plots, respectively. After removing plots without geographic coordinates and all observations on bryophytes and lichens, the database contained 22,195,966 observations on the relative abundance of vascular plant species in a total of 1,117,369 plots. The temporal extent of the data spans from 1885 to 2015, but >95% of vegetation plots were recorded later than 1980. Plot size was reported in 65.4% of plots. While forest plots had plot sizes ≥100 m^2, and in most cases ≤1,000 m^2, non-forest plots typically ranged from 5 to 100 m^2.

Taxonomy. To standardize the nomenclature of species within and between sPlot and TRY (see below), we constructed a taxonomic backbone of the 121,861 names contained in the two databases. Prior to name matching, we ran a series of string manipulation routines in R, to remove special characters and numbers, as well as standardized abbreviations in names. Taxon names were parsed and resolved using Taxonomic Name Resolution Service version 4.0 (TNRS[32]; http://tnrs.iplantcollaborative.org; accessed 20 September 2015), selecting the best match across the five following sources: i) The Plant List (version 1.1; http://www.theplantlist.org/; accessed 19 August 2015), ii) Global Compositae Checklist (GCC, http://compositae.landcareresearch.co.nz/Default.aspx; accessed 21 August 2015), iii) International Legume Database and Information Service (ILDIS, http://www.ildis.org/LegumeWeb; accessed 21 August 2015), iv) Tropicos (http://www.tropicos.org/; accessed 19 December 2014), and v) USDA Plants Database (http://usda.gov/wps/portal/usda/usdahome; accessed 17 January 2015). We allowed for partial matching to the next highest taxonomic rank (genus or family) in cases where full taxon names could not be found. All names matched or converted from a synonym by TNRS were considered accepted taxon names. In cases where no exact match was found (for example when alternative spelling corrections were reported), names with probabilities of ≥ 95% or higher were accepted and those with <95% were examined individually. Remaining non-matching names were resolved using the National Center for Biotechnology Information's Taxonomy database (NCBI, http://www.ncbi.nlm.nih.gov/; accessed 25 October 2011) within TNRS, or sequentially compared directly against The Plant List and Tropicos (accessed September 2015). Names that could not be resolved against any of these lists were left as blanks in the final standardized name field. This resulted in a total of 86,760 resolved names, corresponding to 664 families, occurring in sPlot or TRY or both. Classification into families was carried out according to APGIII[33], and was used to identify non-vascular plant species (~5.1% of the taxon names), which were excluded from the subsequent statistical analysis.

Trait data. Data for 18 traits that are ecologically relevant (Table 1) and sufficiently covered across species[34] were requested from TRY[35] (version 3.0) on 10 August 2016. We applied gap-filling with Bayesian Hierarchical Probabilistic Matrix Factorization (BHPMF[34,36,37]). We used the prediction uncertainties provided by BHPMF for each imputation to assess the quality of gap-filling, and removed all imputations with a coefficient of variation >1[37]. We obtained 18 gap-filled traits for 26,632 out of 58,065 taxa in sPlot, which corresponds to 45.9% of all species but 88.7% of all species × plot combinations. Trait coverage of the most frequent species was 77.2% and 96.2% for taxa that occurred in more than 100 or 1,000 plots, respectively. The gap-filled trait data comprised observed and imputed values on 632,938 individual plants, which we $\log_e$ transformed and aggregated by taxon. For those taxa that were recorded at the genus level only, we calculated genus means. Of 22,195,966 records of vascular plant species with geographic reference, 21,172,989 (95.4%) refer to taxa for which we had gap-filled trait values. This resulted in 1,115,785 and 1,099,463 plots for which we had at least one taxon or two taxa with a trait value (99.5% and 98.1%, respectively, of all 1,121,244 plots), and for which trait means and variances could be calculated.

As some mean values of traits in TRY were based on a very small number of replicates per species, which results in uncertainty in trait mean and variance calculations[38], we tested to what degree the trait patterns in the dataset might be caused by a potential removal of trait variation by imputation of trait values, and we additionally carried out all analyses using the original trait data on the same 632,938 individual plants instead of gap-filled data (Supplementary Table 1). The degree of trait coverage of species ranged between 7.0% and 58.0% for leaf fresh mass and plant height, respectively. Across all species, mean coverage of species with original trait values was 21.8%, compared with 45.9% for gap-filled trait data. Linking these trait values to the species occurrence data resulted in a coverage of species × plot observations with trait values between 7.6% and 96.6% for conduit element length and plant height, respectively (Supplementary Table 1), with a mean of 60.7% compared with 88.7% for those based on gap-filled traits. Using these original trait values to calculate CWM trait values (see below) resulted in a plot coverage of trait values between 48.2% and 100% for conduit element length and SLA, respectively. Across all plots, mean coverage of plots with original trait values was 89.3%, compared with 100% for gap-filled trait data (Supplementary Table 1).

We are aware that using species mean values for traits excludes the possibility of accounting for intraspecific variance, which can also strongly respond to the environment[39]. Thus, using one single value for a species is a source of error in calculating trait means and variances.

Environmental data. We compiled 30 environmental variables (Supplementary Table 2). Macroclimate variables were extracted from CHELSA[40,41], V1.1 (Climatologies at High Resolution for the Earth's Land Surface Areas, www.chelsa-climate.org). CHELSA provides 19 bioclimatic variables equivalent to those used in WorldClim (www.worldclim.org) at a resolution of 30 arcsec (~1 km at the equator), averaging global climatic data from the period 1979–2013 and using a quasi-mechanistic statistical downscaling of the ERA-Interim reanalysis[42].

Variables reflecting growing-season warmth were growing degree days above 1 °C (GDD1) and 5 °C (GDD5), calculated from CHELSA data[43]. We also compiled an index of aridity (AR) and a model for PET extracted from the Consortium of Spatial Information (CGIAR-CSI) website (www.cgiar-csi.org). In addition, 7 soil variables were extracted from the SOILGRIDS project (https://soilgrids.org/, licensed by ISRIC—World Soil Information), downloaded at 250 m resolution and then resampled using the 30 arcsec grid of CHELSA (Supplementary Table 2). We refer to these climate and soil data as 'environmental data'.

Community trait composition. For every trait *j* and plot *k*, we calculated the plot-level trait means as community-weighted mean according to[2,44]:

$$\mathrm{CWM}_{j,k} = \sum_{i}^{n_k} p_{i,k}\, t_{i,j}$$

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

$$\mathrm{CWV}_{j,k} = \sum_{i}^{n_k} p_{i,k}\,(t_{i,j} - \mathrm{CWM}_{j,k})^2$$

CWV is equal to functional dispersion as described by Rao's quadratic entropy (FD_Q)[46], when using a squared Euclidean distance matrix $d_{i,j,k}$[47]:

$$\mathrm{CWV}_{j,k} = \sum_{i}^{n_k} p_{i,k}\,(t_{i,j} - \mathrm{CWM}_{j,k})^2 = \mathrm{FD}_Q = \sum_{i=1}^{n_k-1} \sum_{j=i+1}^{n_k} p_{i,k} p_{j,k} d^2_{i,j,k}$$

We had CWV information for 18 traits for 1,099,463 plots, as at least two taxa were needed to calculate CWV. We performed the calculations using the 'data.table' package[48] in R.

Assessing the degree of filtering. To analyse how plot-level trait means and within-plot trait variances (based on gap-filled trait data) depart from random expectation, for each trait we calculated SESs for the variance in CWMs and for the mean in CWVs. Significantly positive SESs in variance of CWM and significantly negative ones in the mean of CWV can be considered a global-level measure of environmental or biotic filtering. To provide an indication of the global direction of filtering, we also report SESs for the mean of CWM trait values. Similarly,

to measure how much within-community trait dispersion varied globally, we also calculated SESs for the variance in CWV.

We obtained SESs from 100 runs of randomizing trait values across all species globally. In every run we calculated CWM and CWV with random trait values, but keeping all species abundances in plots. Thus, the results of randomization are independent from species co-occurrences structure of plots[49]. For every trait, the SESs of the variance in CWM were calculated as the observed value of variance in CWM minus the mean variance in CWM of the random runs, divided by the standard deviation of the variance in CWM of the random runs (Fig. 1). The SESs for the mean in CWM, the mean in CWV and the variance in CWV were calculated accordingly. Tests for significance of SESs were obtained by fitting generalized Pareto-distribution of the most extreme random values and then estimating *P* values from this fitted distribution[50].

Vegetation trait–environment relationships. Of the 1,115,785 plots with CWM values, 1,114,304 (99.9%) had complete environmental information and coordinates. This set of plots was used to calculate single linear regressions of each of the 18 traits on each of the 30 environmental variables. We used the 'corrplot' function[51] in R to illustrate Pearson correlation coefficients (see Supplementary Figs. 1, 2, 4, 6 and 8) and, for the strongest relationships, produced bivariate graphs and mapped the global distribution of the CWM values using kriging interpolation in ArcGIS 10.2 (Fig. 4). We also tested for non-linear relationships with environment by including an additional quadratic term in the linear model and then reported coefficients of determination. As in the linear relationships of CWM with environment, the highest r^2 values in models with an additional quadratic term were encountered between stem specific density and PET ($r^2=0.156$) and leaf N/P ratio and growing degree days above 5 °C (GDD5, $r^2=0.118$). These were not substantially different from the linear CWM-environment relationships, which had $r^2=0.156$ and $r^2=0.115$, respectively (Fig. 4 and Supplementary Fig. 2). Similarly, including a quadratic term in the regressions did not increase the CWV-environment correlations. Here, the strongest correlations were encountered between plant height and soil pH ($r^2=0.044$) and between SLA and the volumetric content of coarse fragments in the soil (CoarseFrags, $r^2=0.037$), which were similar to those in the linear regressions ($r^2=0.029$ and $r^2=0.036$, respectively, Supplementary Fig. 3).

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

The CWMs and CWVs were scaled to a mean of 0 and s.d. of 1 and then subjected to a PCA, calculated with the 'rda' function from the 'vegan' package[52]. Climate and soil variables were fitted post hoc to the ordination scores of plots of the first two axes, producing correlation vectors using the 'envfit' function. We refrain from presenting any inference statistics, as with >1.1 million plots all environmental variables showed statistically significant correlations. Instead, we report coefficients of determination (r^2), obtained from RDA, using all 30 environmental variables as a constraining matrix, resulting in a maximum of 18 constrained axes corresponding to the 18 traits. We report both r^2 values of the first two axes explained by environment, which is the maximum correlation of the best linear combination of environmental variables to explain the CWM or CWV plot × trait matrix and r^2 values of all 18 constrained axes explained by environment. We plotted the PCA results using the 'ordiplot' function and coloured the points according to the logarithm of the number of plots that fell into grid cells of 0.002 in PCA units (resulting in approximately 100,000 cells). For further details, see the captions of the figures.

Additionally, we carried out the PCA and RDA analyses using CWMs based on original trait values (see above). Because of a poor coverage of some traits we confined the analyses with original trait values to the 6 traits used by Díaz et al.[5], which were leaf area, specific leaf area, leaf N, seed mass, plant height and stem specific density. Using these 6 traits resulted in 954,459 plots that had at least one species with a trait value for each of the 6 traits.

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

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

Resampling procedure in environmental space. To achieve a more even representation of plots across the global climate space, we first subjected the same 30 global climate and soil variables, as described above, to a PCA using the climate space of the whole globe, irrespective of the presence of plots in this space, and scaling each variable to a mean of zero and a standard deviation of one. We used a 2.5 arcmin spatial grid, which comprised 8,384,404 terrestrial grid cells. We then counted the number of vegetation plots in the sPlot database that fell into each grid cell. For this analysis, we did not use the full set of 1,117,369 plots with trait information (see above), but only those plots that had a location inaccuracy of max. 3 km, resulting in a total of 799,400 plots. The resulting PCA scores based on the first two principal components (PC1–PC2) were rasterized to a 100 × 100 grid in PC1–PC2 environmental space, which was the most appropriate resolution according to a sensitivity analysis. This sensitivity analysis tested different grid resolutions, from a coarse-resolution bivariate space of 100 grid cells (10 × 10) to a very fine-resolution space of 250,000 grid cells (500 × 500), iteratively increasing the number of cells along each principal component by 10 cells. For each iteration, we computed the total number of plots from sPlot per environmental grid cell and plotted the median sampling effort (number of plots) across all grid cells versus the resolution of the PC1–PC2 space. We found that the curve flattens off at a bivariate environmental space of 100 × 100 grid cells, which was the resolution for which the median sampling effort stabilized at around 50 plots per grid cell. As a result, we resampled plots only in environmental cells with more than 50 plots (858 cells in total).

To optimize our resampling procedure within each grid cell, we used the heterogeneity-constrained random (HCR) resampling approach[53]. The HCR approach selects the subset of vegetation plots for which those plots are the most dissimilar in their species composition while avoiding selection of plots representing peculiar and rare communities that differ markedly from the main set of plant communities (outliers), thus providing a representative subset of plots from the resampled grid cell. We used the turnover component of the Jaccard's dissimilarity index (β_{jtu}[54]) as a measure of dissimilarity. The β_{jtu} index accounts for species replacement without being influenced by differences in species richness. Thus, it reduces the effects of any imbalances that may exist between different plots due to species richness. We applied the HCR approach within a given grid cell by running 1,000 iterations, randomly selecting 50 plots out of the total number of plots available within that grid cell. Where the cell contained 50 or fewer plots, all were included and the resampling procedure was not run. This procedure thinned out over-sampled climate types, while retaining the full environmental gradient.

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

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

The mean PCA loadings across these 100 subsets (summarized in Supplementary Fig. 10) were fully consistent with those of the full dataset in Fig. 2, with the same two functional continua in plant size and diaspore mass (from bottom left to top right), and perpendicular to that, the leaf economics spectrum. The variation in CWM accounted for by the first two axes was on average 50.9% ± 0.04 s.d., and thus virtually identical with that in the total dataset. In contrast, the variation explained on average by macroclimate and soil conditions (26.5% ± 0.01 s.d. as the average cumulative variance of all 18 constrained axes in the RDAs across all 100 runs) was considerably larger than that for the total dataset, which is also reflected in consistently higher correlations between traits and environmental variables (Supplementary Fig. 11). The highest mean correlation was encountered for plant height and PET (mean $r^2 = 0.342$ across 100 runs). PET was a better predictor for plant height than the precipitation of the wettest months (bio13, mean $r^2 = 0.231$), as had been suggested previously[6]. The correlation of PET with stem specific density (mean $r^2 = 0.284$) and warmth of the growing season (expressed as growing degree days above the threshold 5 °C, GDD5) with leaf N/P ratio (mean $r^2 = 0.250$) ranked among the best 12 correlations encountered out of all 540 trait–environment relationships, which confirms the patterns found in the whole dataset (compared with Fig. 4). Overall, the coefficients of determination were much closer to the ones reported from other studies with a global collection of a few hundred plots (r^2 values ranging from 36% to 53% based on multiple regressions of single traits with 5 to 6 environmental drivers[22]).

Reporting Summary. Further information on experimental design is available in the Nature Research Reporting Summary linked to this article.

Data availability

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

Received: 2 March 2018; Accepted: 18 September 2018;
Published online: 19 November 2018

References

1. Warming, E. *Lehrbuch der ökologischen Pflanzengeographie – Eine Einführung in die Kenntnis der Pflanzenvereine* (Borntraeger, Berlin, 1896).
2. Garnier, E. et al. Plant functional markers capture ecosystem properties during secondary succession. *Ecology* **85**, 2630–2637 (2004).
3. Ordoñez, J. C. A global study of relationships between leaf traits, climate and soil measures of nutrient fertility. *Glob. Ecol. Biogeogr.* **18**, 137–149 (2009).
4. Garnier, E., Navas, M.-L. & Grigulis, K. *Plant Functional Diversity – Organism Traits, Community Structure, and Ecosystem Properties* (Oxford Univ. Press, Oxford, 2016).
5. Díaz, S. et al. The global spectrum of plant form and function. *Nature* **529**, 167–171 (2016).
6. Moles, A. T. et al. Global patterns in plant height. *J. Ecol.* **97**, 923–932 (2009).
7. Wright, I. J. et al. The worldwide leaf economics spectrum. *Nature* **428**, 821–827 (2004).
8. Reich, P. B. The world-wide 'fast–slow' plant economics spectrum: a traits manifesto. *J. Ecol.* **102**, 275–301 (2014).
9. Adler, P. B. et al. Functional traits explain variation in plant life history strategies. *Proc. Natl Acad. Sci. USA* **111**, 740–745 (2014).
10. Marks, C. O. & Lechowicz, M. J. Alternative designs and the evolution of functional diversity. *Am. Nat.* **167**, 55–67 (2006).
11. Grime, J. P. Trait convergence and trait divergence in herbaceous plant communities: mechanisms and consequences. *J. Veg. Sci.* **17**, 255–260 (2006).
12. Muscarella, R. & Uriarte, M. Do community-weighted mean functional traits reflect optimal strategies?. *Proc. R. Soc. B* **283**, 20152434 (2016).
13. Swenson, N. G. & Weiser, M. D. Plant geography upon the basis of functional traits: an example from eastern North American trees. *Ecology* **91**, 2234–2241 (2010).
14. Fyllas, N. M. et al. Basin-wide variations in foliar properties of Amazonian forest: phylogeny, soils and climate. *Biogeosciences* **6**, 2677–2708 (2009).
15. Swenson, N. G. et al. Phylogeny and the prediction of tree functional diversity across novel continental settings. *Glob. Ecol. Biogeogr.* **26**, 553–562 (2017).
16. Swenson, N. G. et al. The biogeography and filtering of woody plant functional diversity in North and South America. *Glob. Ecol. Biogeogr.* **21**, 798–808 (2012).
17. Wright, I. J. et al. Global climatic drivers of leaf size. *Science* **357**, 917–921 (2017).
18. Mayfield, M. M. & Levine, J. M. Opposing effects of competitive exclusion on the phylogenetic structure of communities. *Ecol. Lett.* **13**, 1085–1093 (2010).
19. Kraft, N. J. B. et al. Community assembly, coexistence and the environmental filtering metaphor. *Funct. Ecol.* **29**, 592–599 (2015).
20. Barboni, D. et al. Relationships between plant traits and climate in the Mediterranean region: a pollen data analysis. *J. Veg. Sci* **15**, 635–646 (2004).
21. Borgy, B. et al. Plant community structure and nitrogen inputs modulate the climate signal on leaf traits. *Glob. Ecol. Biogeogr.* **26**, 1138–1152 (2017).
22. van Bodegom, P. M., Douma, J. C. & Verheijen, L. M. A fully traits-based approach to modeling global vegetation distribution. *Proc. Natl Acad. Sci. USA* **111**, 13733–13738 (2014).
23. Moles, A. T. et al. Which is a better predictor of plant traits: temperature or precipitation? *J. Veg. Sci.* **25**, 1167–1180 (2014).
24. Ordoñez, J. C. et al. Plant strategies in relation to resource supply in mesic to wet environments: does theory mirror nature?. *Am. Nat.* **175**, 225–239 (2010).
25. Simpson, A. J., Richardson, S. J. & Laughlin, D. C. Soil–climate interactions explain variation in foliar, stem, root and reproductive traits across temperate forests. *Glob. Ecol. Biogeogr.* **25**, 964–978 (2016).
26. Lienin, P. & Kleyer, M. Plant leaf economics and reproductive investment are responsive to gradients of land use intensity. *Agric. Ecosyst. Environ.* **145**, 67–76 (2011).
27. Maire, V. et al. Habitat filtering and niche differentiation jointly explain species relative abundance within grassland communities along fertility and disturbance gradients. *New Phytol.* **196**, 497–509 (2012).
28. Craine, J. M. et al. Global patterns of foliar nitrogen isotopes and their relationships with climate, mycorrhizal fungi, foliar nutrient concentrations, and nitrogen availability. *New Phytol.* **183**, 980–992 (2009).
29. Güsewell, S. N:P ratios in terrestrial plants: variation and functional significance. *New Phytol.* **164**, 243–266 (2004).
30. Reich, P. B. & Oleksyn, J. Global patterns of plant leaf N and P in relation to temperature and latitude. *Proc. Natl Acad. Sci. USA* **101**, 11001–11006 (2004).
31. Scheiter, S., Langan, L. & Higgins, S. I. Next generation dynamic global vegetation models: learning from community ecology. *New Phytol.* **198**, 957–969 (2013).
32. Boyle, B. et al. The Taxonomic Name Resolution Service: an online tool for automated standardization of plant names. *BMC Bioinform.* **14**, 16 (2013).
33. Bremer, B. et al. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III. *Bot. J. Linn. Soc.* **161**, 105–121 (2009).
34. Schrodt, F. et al. BHPMF – a hierarchical Bayesian approach to gap-filling and trait prediction for macroecology and functional biogeography. *Glob. Ecol. Biogeogr.* **24**, 1510–1521 (2015).
35. Kattge, J. et al. TRY – a global database of plant traits. *Glob. Change Biol.* **17**, 2905–2935 (2011).
36. Shan, H. et al. Gap filling in the plant kingdom – trait prediction using hierarchical probabilistic matrix factorization. In *Proc. 29th Int. Conf. Machine Learning (ICML 2012)* 1303–1310 (Omnipress, Madison, 2012).
37. Fazayeli, F. et al. Uncertainty quantified matrix completion using Bayesian Hierarchical Matrix factorization. In *Proc. 13th Int. Conf. Machine Learning and Applications (ICMLA 2014)* 312–317 (Institute of Electrical and Electronics Engineers, Danvers, 2014).
38. Borgy, B. et al. Sensitivity of community-level trait–environment relationships to data representativeness: a test for functional biogeography. *Glob. Ecol. Biogeogr.* **26**, 729–739 (2017).
39. Herz, K. et al. Drivers of intraspecific trait variation of grass and forb species in German meadows and pastures. *J. Veg. Sci.* **28**, 705–716 (2017).
40. Karger, D. N. et al. Climatologies at high resolution for the Earth's land surface areas. *Sci. Data* **4**, 170122 (2017).
41. Karger, D. N. et al. *Climatologies at High Resolution for the Earth Land Surface Areas (Version 1.1)* (World Data Center for Climate (WDCC) at DKRZ, 2016); http://chelsa-climate.org/downloads/
42. Dee, D. P. et al. The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Q. J. R. Meteorol. Soc* **137**, 553–597 (2011).
43. Synes, N. W. & Osborne, P. E. Choice of predictor variables as a source of uncertainty in continental-scale species distribution modelling under climate change. *Glob. Ecol. Biogeogr.* **20**, 904–914 (2011).

44. Enquist, B. et al. Scaling from traits to ecosystems: developing a general trait driver theory via integrating trait-based and metabolic scaling theories. *Adv. Ecol. Res.* **52**, 249–318 (2015).
45. Buzzard, V. et al. Re-growing a tropical dry forest: functional plant trait composition and community assembly during succession. *Funct. Ecol.* **30**, 1006–1013 (2016).
46. Rao, C. R. Diversity and dissimilarity coefficients: a unified approach. *Theor. Popul. Biol.* **21**, 24–43 (1982).
47. Champely, S. & Chessel, D. Measuring biological diversity using Euclidean metrics. *Environ. Ecol. Stat* **9**, 167–177 (2002).
48. Dowle, M. et al. *data.table: Extension of data.frame*. R Package Version 1.9.6 (2015); https://CRAN.R-project.org/package=data.table
49. Hawkins, B. A. et al. Structural bias in aggregated species-level variables driven by repeated species co-occurrences: a pervasive problem in community and assemblage data. *J. Biogeogr.* **44**, 1199–1211 (2017).
50. Knijnenburg, T. A. et al. Fewer permutations, more accurate P-values. *Bioinformatics* **25**, i161–i168 (2009).
51. Friendly, M. Corrgrams: exploratory displays for correlation matrices. *Am. Stat.* **56**, 316–324 (2002).
52. Oksanen, J. et al. *vegan: Community Ecology Package*. R Package Version 2.3-3 (2016); https://CRAN.R-project.org/package=vegan
53. Lengyel, A., Chytrý, M. & Tichý, L. Heterogeneity-constrained random resampling of phytosociological databases. *J. Veg. Sci* **22**, 175–183 (2011).
54. Baselga, A. The relationship between species replacement, dissimilarity derived from nestedness, and nestedness. *Glob. Ecol. Biogeogr.* **21**, 1223–1232 (2012).
55. Garnier, E. et al. Towards a thesaurus of plant characteristics: an ecological contribution. *J. Ecol.* **105**, 298–309 (2017).

Acknowledgements

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

Author contributions

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

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41559-018-0699-8.

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

Correspondence and requests for materials should be addressed to H.B.

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

[1]Martin Luther University Halle-Wittenberg, Institute of Biology/Geobotany and Botanical Garden, Halle, Germany. [2]German Centre for Integrative Biodiversity Research (iDiv), Halle-Jena-Leipzig, Leipzig, Germany. [3]Zurich University of Applied Sciences, Institute of Natural Resource Sciences, Research Group Vegetation Ecology, Wädenswil, Switzerland. [4]University of Bayreuth, Bayreuth Center of Ecology and Environmental Research, Plant Ecology, Bayreuth, Germany. [5]CNRS, Université de Picardie Jules Verne, UR 'Ecologie et Dynamique des Systèmes Anthropisés' (EDYSAN, UMR 7058 CNRS-UPJV), Amiens, France. [6]Research Unit of Biodiversity (CSIC/UO/PA), University of Oviedo, Campus de Mieres, Mieres, Spain. [7]Wageningen Environmental Research (Alterra), Team Vegetation, Forest and Landscape Ecology, Wageningen, The Netherlands. [8]MTA Centre for Ecological Research, GINOP Sustainable Ecosystems Group, Tihany, Hungary. [9]Masaryk University, Department of Botany and Zoology, Brno, Czech Republic. [10]University of Nottingham, School of Geography, University Park, Nottingham, UK. [11]University of Rostock, Faculty for Agricultural and Environmental Sciences, Rostock, Germany. [12]Max Planck Institute for Biogeochemistry, Jena, Germany. [13]Universidade Federal do Rio Grande do Sul, Department of Ecology, Porto Alegre, Brazil. [14]University of North Carolina at Chapel Hill, Department of Biology, Chapel Hill, NC, USA. [15]Santa Clara University, Department of Biology, Santa Clara, CA, USA. [16]Leiden University, Institute of Environmental Sciences, Department Conservation Biology, Leiden, The Netherlands. [17]Institute of Botany of the Czech Academy of Sciences, Průhonice, Czech Republic. [18]Escuela de Ciencias Agropecuarias y Ambientales – ECAPMA, Universidad Nacional Abierta y a Distancia – UNAD, Sede José Celestino Mutis, Bogotá, Colombia. [19]Shahjalal University of Science and Technology, Department of Forestry and Environmental Science, Sylhet, Bangladesh. [20]University of Bayreuth, Bayreuth Center of Ecology and Environmental Research, Department of Disturbance Ecology, Bayreuth, Germany. [21]Sapienza University of Rome, Department of Environmental Biology, Rome, Italy. [22]Great Lakes Forestry Centre, Canadian Forest Service, Natural Resources Canada, Sault Ste Marie, Ontario, Canada. [23]Florida International University, Department of Biological Sciences, International Center for Tropical Botany, Miami, FL, USA. [24]Universidade Federal do Acre, Campus de Cruzeiro do Sul, Acre, Brazil. [25]Ghent University, Faculty of Bioscience Engineering, Department of Green Chemistry and Technology (ISOFYS) and Department of Environment (CAVELab), Gent, Belgium. [26]University of Göttingen, Albrecht von Haller Institute of Plant Sciences, Vegetation Analysis & Plant Diversity, Göttingen, Germany. [27]University of the Basque Country UPV/EHU, Bilbao, Spain. [28]Aarhus University, Department of Bioscience, Biodiversity Dynamics in a Changing World (BIOCHANGE) & Section for Ecoinformatics & Biodiversity, Aarhus, Denmark. [29]University of Oxford, Environmental Change Institute, School of Geography and the Environment, Oxford, UK. [30]Rocky Mountain Biological Laboratory, Crested Butte, CO, USA. [31]Scientific Research Center of the Slovenian Academy of Sciences and Arts, Institute of Biology, Ljubljana, Slovenia. [32]University of Nova Gorica, Nova Gorica, Slovenia. [33]Universidad Rey Juan Carlos, Department of Biology, Geology, Physics and Inorganic Chemistry, Madrid, Spain. [34]Czech University of Life Sciences, Faculty of Forestry and Wood Science, Department of Forest Ecology, Prague, Czech Republic. [35]Vrije Universiteit Amsterdam, Faculty of Science, Department of Ecological Science, Amsterdam, The Netherlands. [36]Helmholtz Centre for Environmental Research – UFZ, Department of Community Ecology, Halle, Germany. [37]University of Würzburg, Department of Animal Ecology and Tropical Biology, Würzburg, Germany. [38]Cirad, UMR EcoFoG, Campus Agronomique, Kourou, French Guiana. [39]Instituto Multidisciplinario de Biología Vegetal, CONICET and FCEFyN, Universidad Nacional de Córdoba, Córdoba, Argentina. [40]Wake Forest University, Department of Biology, Winston Salem, NC, USA. [41]Universidad Nacional de San Antonio Abad del Cusco, Herbario Vargas (CUZ), Cusco, Peru. [42]University of Exeter, College of Life and Environmental Sciences, Geography, Exeter, UK. [43]Université du Québec en Abitibi-Témiscamingue, Institut de recherche sur les forêts, Rouyn-Noranda, Quebec, Canada. [44]CNRS, Université de Montpellier, Université Paul-Valéry Montpellier, EPHE, Centre d'Ecologie Fonctionnelle et Evolutive (UMR5175), Montpellier, France. [45]University of Adelaide, Terrestrial Ecosystem Research Network, School of Biological Sciences, Adelaide, South Australia, Australia. [46]Universidad de Chile, Facultad de Ciencias Agronómicas, Departamento de Ciencias Ambientales y Recursos Naturales Renovables, Santiago, Chile. [47]Institut Français de Recherche pour l'Exploitation de la MER, UMR 248 MARBEC (CNRS, IFREMER, IRD, UM), Sète, France. [48]University of British Columbia, The Department of Geography, Vancouver, British Columbia, Canada. [49]Institut National Polytechnique Félix Houphouët-Boigny, Yamoussoukro, Côte d'Ivoire. [50]Cirad, University Montpellier, UR Forests & Societies, Montpellier, France. [51]Universidade do Estado de Santa Catarina, Departamento de Engenharia Florestal, Lages, Brazil. [52]University of Münster, Institute of Landscape Ecology, Münster, Germany.

[53]University of Göttingen, Plant Ecology and Ecosystems Research, Göttingen, Germany. [54]University of Hamburg, Biodiversity, Biocenter Klein Flottbek and Botanical Garden, Hamburg, Germany. [55]University of Wroclaw, Institute of Environmental Biology, Department of Vegetation Ecology, Wrocław, Poland. [56]University of Zurich, Department of Systematic and Evolutionary Botany, Zurich, Switzerland. [57]Swiss Federal Research Institute WSL, Birmensdorf, Switzerland. [58]University of Oldenburg, Institute of Biology and Environmental Sciences, Landscape Ecology Group, Oldenburg, Germany. [59]Central Siberian Botanical Garden SB RAS, Novosibirsk, Russia. [60]University of Waikato, Environmental Research Institute, School of Science, Hamilton, New Zealand. [61]University of Wyoming, Department of Botany, Laramie, WY, USA. [62]Leiden University, Naturalis Biodiversity Center, Leiden, The Netherlands. [63]Agroecology, University of Göttingen, Göttingen, Germany. [64]UCA, INRA, VetAgro Sup, UREP, Clermont-Ferrand, France. [65]University of Sofia, Faculty of Biology, Department of Ecology and Environmental Protection, Sofia, Bulgaria. [66]University of Oxford, Environmental Change Institute, School of Geography and the Environment, Oxford, UK. [67]ICREA, Barcelona, Spain. [68]CREAF, Barcelona, Spain. [69]Royal Botanic Gardens Kew, Millennium Seed Bank, Conservation Science, Ardingly, UK. [70]AMAP, IRD, CIRAD, CNRS, INRA, Université Montpellier, Montpellier, France. [71]University of Edinburgh, School of GeoSciences, Edinburgh, UK. [72]Universidad Estatal Amazónica, Conservación y Manejo de Vida Silvestre, Puyo, Ecuador. [73]Estonian University of Life Science, Department of Crop Science and Plant Biology, Tartu, Estonia. [74]Landcare Research, Lincoln, New Zealand. [75]Radboud University Nijmegen, Institute for Water and Wetland Research, Nijmegen, The Netherlands. [76]CSIC, Global Ecology Unit, CREAF-CEAB-UAB, Cerdanyola del Vallès, Spain. [77]University of Barcelona, Faculty of Biology, Department of Evolutionary Biology, Ecology and Environmental Sciences, Barcelona, Spain. [78]Center for Advanced Studies of Blanes, Spanish Research Council (CEAB-CSIC), Blanes, Spain. [79]University of Leeds, School of Geography, Leeds, UK. [80]University of Tartu, Tartu, Estonia. [81]University of Minnesota, Department of Forest Recourses, St. Paul, MN, USA. [82]Western Sydney University, Hawkesbury Institute for the Environment, Sydney, New South Wales, Australia. [83]Friedrich Schiller University Jena, Institute of Ecology and Evolution, Jena, Germany. [84]Universidade Regional de Blumenau, Departamento de Engenharia Florestal, Blumenau, Brazil. [85]Senckenberg Biodiversity and Climate Research Centre (BiK-F), Data and Modelling Centre, Frankfurt am Main, Germany. [86]University of La Serena, Department of Biology, La Serena, Chile. [87]Universidade Federal do Acre, Museu Universitário / Centro de Ciências Biológicas e da Natureza / Laboratório de Botânica e Ecologia Vegetal, Rio Branco, Brazil. [88]Peking University, College of Urban and Environmental Sciences, Beijing, China. [89]Iwokrama International Centre for Rain Forest Conservation and Development, Georgetown, Guyana. [90]Aristotle University of Thessaloniki, School of Biology, Department of Botany, Thessaloniki, Greece. [91]Bulgarian Academy of Sciences, Institute of Biodiversity and Ecosystem Research, Sofia, Bulgaria. [92]Helmholtz Center for Environmental Research – UFZ, Department of Physiological Diversity, Leipzig, Germany. [93]University of Oulu, Department of Ecology & Genetics, Oulu, Finland. [94]University of Wisconsin – Eau Claire, Department of Biology, Eau Claire, WI, USA. [95]Senckenberg Museum of Natural History Görlitz, Görlitz, Germany. [96]TU Dresden, International Institute (IHI) Zittau, Zittau, Germany. [97]University of Leipzig, Systematic Botany and Functional Biodiversity, Leipzig, Germany. *e-mail: helge.bruelheide@botanik.uni-halle.de