Global trade-offs in tree functional traits

Daniel S. Maynard[1*], Lalasia Bialic-Murphy[1], Constantin M. Zohner[1], Colin Averill[1], Johan van den Hoogen[1], Haozhi Ma[1], Lidong Mo[1], Gabriel Reuben Smith[1,2], Isabelle Aubin[3], Erika Berenguer[4,5], Coline C.F. Boonman[6], Jane Catford[7], Bruno E. L. Cerabolini[8], Arildo S. Dias[9], Andrés González-Melo[10], Peter Hietz[11], Christopher H. Lusk[12], Akira S. Mori[13], Ülo Niinemets[14], Valério D. Pillar[15], Julieta A. Rosell[16], Frank M. Schurr[17], Serge N. Sheremetev[18], Ana Carolina da Silva[19], Ênio Sosinski[20], Peter M. van Bodegom[21], Evan Weiher[22], Gerhard Bönisch[23], Jens Kattge[24,25], Thomas W. Crowther[1]

**Corresponding author email:* dmaynard@ethz.ch

Affiliations

(1) Institute of Integrative Biology, ETH Zürich, 8092 Zürich, Switzerland; (2.) Department of Biology, Stanford University, Stanford, CA 94305, USA; (3) Natural Resources Canada, Canadian Forest Service, Great Lakes Forestry Centre, Sault Ste Marie, ON P6A 2E5, Canada; (4) Environmental Change Institute, School of Geography and the Environment, University of Oxford, Oxford, UK; (5) Lancaster Environment Centre, Lancaster University, Lancaster, UK; (6) Department of Aquatic Ecology & Environmental Biology, Institute for Water and Wetland Research, Radboud University, Nijmegen, The Netherlands; (7) Department of Geography, King's College London, 30 Aldwych, London, WC2B 4BG, UK; (8) Department of Biotechnologies and Life Sciences (DBSV), University of Insubria, 21100 Varese, Italy; (9) Goethe University, Institute for Physical Geography, Altenhöferallee 1, 60438 Frankfurt am Main, Germany; (10) Biology Department, Faculty of Natural Sciences, Universidad del Rosario. Avenida carrera 24 # 63C-69. Bogotá, Colombia; (11) Institute of Botany, University of Natural Resources and Life Sciences, Gregor Mendel St. 33, 1190 Vienna, Austria; (12) Environmental Research Institute, University of Waikato, Hamilton, New Zealand; (13) Faculty of Environment and Information Sciences, Yokohama National University, 79-7 Tokiwadai, Hodogaya, Yokohama, 240-8501, Japan; (14) Chair of Crop Science and Plant Biology, Estonian University of Life Sciences, Tartu 51006, Estonia; (15) Department of Ecology, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, 91501-970, Brazil.; (16) Laboratorio Nacional de Ciencias de la Sostenibilidad, Instituto de Ecología, Universidad Nacional Autónoma de México, A.P. 70-275, Ciudad Universitaria, Coyoacán, 04510 Mexico City, Mexico; (17) Institute of Landscape and Plant Ecology, University of Hohenheim, Ottilie-Zeller-Weg 2, D-70599 Stuttgart, Germany; (18) Komarov Botanical Institute, Prof. Popov str., 2, St. Petersburg, 197376, Russia; (19) Department of Forestry, Santa Catarina State University, 88520-000, Lages, SC, Brazil; (20) Embrapa Clima Temperado, Pelotas, RS, Brazil, 96010-971; (21) Institute of Environmental Science, Leiden University, 2333 CC Leiden, the Netherlands; (22) Department of Biology, University of Wisconsin – Eau Claire, Eau Claire, WI 54702 USA; (23) Max Planck Institute for Biogeochemistry, 07745 Jena, Germany; (24) German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Deutscher Platz 5e, 04103 Leipzig, Germany

ABSTRACT

Due to massive energetic investments in woody support structures, trees are subject to unique physiological, mechanical, and ecological pressures not experienced by herbaceous plants. When considering trait relationships across the entire plant kingdom, plant trait frameworks typically must omit traits unique to large woody species, thereby limiting our understanding of how these distinct ecological pressures shape trait relationships in trees. Here, by considering 18 functional traits—reflecting leaf economics, wood structure, tree size, reproduction, and below-ground allocation—we quantify the major axes of variation governing trait expression of trees worldwide. We show that trait variation within and across angiosperms and gymnosperms is captured by two independent processes: one reflecting tree size and competition for light, the other reflecting leaf photosynthetic capacity and nutrient economies. By exploring multidimensional relationships across clusters of traits, we further identify a representative set of seven traits which captures the majority of variation in form and function in trees: maximum tree height, stem conduit diameter, specific leaf area, seed mass, bark thickness, root depth, and wood density. Collectively, this work informs future trait-based research into the functional biogeography of trees, and contributes to our fundamental understanding of the ecological and evolutionary controls on forest biodiversity and productivity worldwide.

INTRODUCTION

Physiological and morphological traits determine the water, nutrient, and light economies of trees, directly influencing how individuals interact with each other and with the surrounding environment[1–5]. Traits that elevate performance in one habitat typically reduce performance in others, leading to selection for specific traits across environments[6]. Genetic, morphological, and biophysical constraints subsequently limit the range of traits that a species can exhibit, leading to so-called trait 'trade-offs' that shape species' geographic distributions[7], coexistence mechanisms[8,9], and the provision of ecosystem services[10,11]. Despite a wealth of research into trait trade-offs across the plant kingdom, there remains relatively little understanding of the unique trade-offs faced by large woody species. Identifying the dominant trait trade-offs in trees is fundamental to our understanding of the functional biogeography of forests, and critical for predicting how forest diversity, composition, and function will respond to changing environmental conditions[12–15].

Prior studies have identified a key set of traits that summarize the spectrum of form and function across herbaceous and woody plants[4,16–18]. But due to their size, longevity, ontogeny, and unique structural properties, trees have distinct characteristics and face novel abiotic stressors relative to herbaceous plants[5,19–23]. Trait analyses which include both woody and herbaceous plants are forced to omit critical aspects of tree architecture (e.g., bark properties, crown size, stem conduit diameter), which overlooks the massive energetic investments in structures that are unique to large woody species[22,24]. Understanding how tree-specific traits align with existing plant trait frameworks is key for identifying the dominant biogeographic and ecological processes governing forest structure across broad spatial scales[15].

Here, we uses a global database[25] of more than 350,000 trait measurements to explore relationships among 18 functional traits, reflecting leaf economics, wood structure and function, tree size and architecture, reproduction, and below-ground allocation (Fig. 1b). We asked: (1) Which dominant trade-offs and abiotic variables best capture overall variation in tree trait expression? and (2) What are the dominant multi-trait constellations that capture the breadth of form and function at the global scale? We hypothesized that traits related to canopy architecture, wood density, and tree size would emerge as key variables governing trait patterns in trees, driven by the large energetic requirements of large woody structures[26,27]. Collectively, this work identifies emergent constraints

on tree functional biogeography, and sheds light on the core ecological processes shaping functional trait expression in forests worldwide.

RESULTS & DISCUSSION

TRADE-OFFS IN TREE TRAIT EXPRESSION. Our analysis included 386,526 unique trait measurements across 18 traits, encompassing 6905 tree species from 1629 genera and 203 families (Fig. 1). Traits were measured at 8490 distinct locations, including every continent except Antarctica. To explore trade-offs in functional traits at the individual level, we used random forest machine learning models to estimate trait expression for each individual as a function of its environment and phylogenetic history. Environmental predictors included a range of climate[28–31], soil[32], topographic[33], and geological[34] features. Phylogenetic history was incorporated via phylogenetic eigenvectors[35,36] (see Methods).

Across all 18 traits, our models explained 54% of trait variation (buffered leave-one-out cross-validation, see Methods), with a relative predictive error of ± 28% (Fig. S1-S2). The inclusion of environmental variables led to substantial increases in explanatory power and accuracy, reducing the expected predictive error of the models by 10% and improving the explanatory power by 35% across all traits. Overall, environmental variables and phylogenetic information had approximately equal explanatory power (relative importance of 0.51 vs 0.49 for phylogeny vs. environment), albeit with substantial variation across traits (Fig. S3). Traits with high intraspecific variation and ontogenetic plasticity exhibited particularly strong increases in accuracy with the inclusion of environmental variables (e.g., 19% and 16% improvement for crown height and root depth, respectively). Only seed dry mass had no residual environmental signal after accounting for phylogeny (Fig. S2-S3).

Using the resulting trait models, we quantified trait trade-offs at the individual levels, accounting for both phylogenetic trait conservatism and environmental-mediated trait variation. When considering all traits simultaneously, the first two axes of the resulting principal-component analysis capture 40% of variation in overall trait expression (Figs. 2a, S4). The first trait axis correlates most strongly with leaf thickness (ρ = -0.77), specific leaf area (ρ = 0.74), and leaf nitrogen (ρ = 0.70). By capturing key aspects of the leaf economic spectrum[16], these traits reflect various physiological controls on leaf-level resource processing, tissue turnover and

photosynthetic rate[6,37,38]. Thick leaves with low SLA can help minimize desiccation, herbivory, frost damage, and nutrient limitation, but at the cost of reduced photosynthetic potential due to primary investment in structural resistance[39]. Accordingly, leaf nitrogen—a crucial component of Rubisco for photosynthesis[40,41]—trades off strongly with leaf thickness (ρ = -0.47). By reflecting an organismal-level trade-off between photosynthetic potential in optimal conditions and abiotic tolerance in suboptimal conditions, this first axis thus captures the core distinction between "acquisitive" and "conservative" functional traits which underpin fast-slow life-history strategies across the plant kingdom[6,17,18]. At one end of this spectrum are species with acquisitive traits which confer higher growth, faster nutrient cycling, and greater photosynthetic potential. At the other end are conservative traits (thick leaves) which help trees withstand a variety of stressful abiotic conditions, but which come at the cost of reduced photosynthetic capacity and growth in optimal environments.

The second trait axis correlates most strongly with maximum tree height (ρ = 0.72), crown height, (ρ = 0.70), and crown diameter (ρ = 0.82), highlighting the overarching importance of competition for light and canopy position in forest [6] (Figs. 2a, S4). Large trees and large crowns are critical for light access and for maximizing light interception down through the canopy[42]. Nevertheless, tall trees with deep canopies also experience greater susceptibility to disturbance and mechanical damage, primarily due to wind and weight[43,44]. Because of the massive carbon and nutrient costs required to create large woody structures[26,27], larger trees are less viable in nutrient-limited or colder climates[45], and in exposed areas with high winds or extreme weather events[46]. This second axis thus reflects a fundamental biotic/abiotic trade-off related to overall tree size, which is largely orthogonal to leaf-level nutrient-use and photosynthetic capacity.

Despite substantial differences in wood and leaf structures between angiosperms and gymnosperms (e.g., vessels vs. tracheids), the two main trade-offs hold within, as well as across, clades (Fig. 2b-c, S5-S6). Gymnosperms, however, exhibit less orthogonality between these axes, in part due to less variation in leaf thickness and a stronger subsequent correlation between leaf thickness and tree size (ρ = 0.39 vs. 0.04 for gymnosperms vs angiosperms). Nevertheless, despite differences in physiology and morphology, gymnosperms and angiosperms are subject to the same physical, mechanical, and chemical processes that determine the ability to withstand various biotic

and abiotic pressures[7,47,48]. Our results show that these processes translate into similar fundamental constraints on trait expression across clades.

Collectively, the two primary trait axes thus reflect two different aspects of the dominance-tolerance trade-off: (1) the ability to maximize leaf photosynthetic activity, at the cost of increase risk of leaf desiccation, and (2) the ability to compete for space and maximize light interception via tree size, at the cost of increased susceptibility to mechanical damage. Notably, these two trade-offs closely mirror those seen when considering herbaceous species alongside woody species[4], though we observe stronger orthogonality between leaf function and plant size is found when considering the whole plant kingdom. Thus, rather than fundamentally reshape the dominant trade-offs, the inclusion of tree-specific traits shows that these two ecological constraints systematically affect all aspects of tree form and function, illustrating the universality of these two ecological trade-offs across the plant kingdom.

ENVIRONMENTAL PREDICTORS OF TRAIT TRADE-OFFS. To examine how environmental variation shapes trait expression across the globe, we quantified the relationships between environmental conditions and the dominant trait trade-offs.

In line with previous analysis[49], temperature variables were the strongest univariate drivers of trait trade-offs (Figs. 3, S7). The first PC axis (leaf thickness) correlates most strongly with annual mean temperature ($\rho = 0.26$, Fig. 3a), reflecting that leaves face increased frost risk and reduced photosynthetic potential in colder conditions. Thus, selection should favor thick leaves with low SLA over thin leaves with high SLA and high nutrient-use[37]. However, the univariate environmental signal is relatively weak (Fig. 3c), highlighting that the first PC axis captures more complex relationships among leaf-level economies and environmental conditions (e.g. between angiosperms vs. gymnosperms).

The second PC axis (tree size) correlates most strongly with temperature annual range (Fig. 3b). Crown diameter, leaf area, and tree height, in particular, all exhibit strong negative correlations with temperature annual range ($\rho = -0.65$, -0.56 and -0.54, respectively, Fig. 3c). At the global scale, high temperature variation is inversely correlated with annual temperature ($\rho = -0.79$), such that larger structural components are favored in areas with consistently warm temperatures—primarily tropical regions near the equator, and coastal regions throughout the Americas, Australia,

and southern Africa. Trees in such environments are more likely to experience strong biotic interactions, which should increase evolutionary and ecological selection pressures over time[50,51], favoring species with high competitive ability and efficient light acquisition strategies.

Despite the primary importance of temperature governing tree trait trade-offs, precipitation and temperature regimes are highly correlated, and the main climate stressors to trees arise via interactions between temperature and water availability (e.g., xylem cavitation and embolism, fire regimes, and leaf desiccation). Indeed, when exploring the bivariate drivers of trait expression, precipitation variables emerge as the strongest secondary drivers of each trait trade-off. The first PC axis exhibits the highest values at high temperatures in combination with high precipitation (Fig. 3c), reflecting broad-scale differences in habitat requirements across angiosperms and gymnosperms[52,53]. For the second PC axis, higher values are observed among trees in regions with low temperature variation in tandem with sufficient ground water access (Fig. 3d), demonstrating that soil hydrology and precipitation place key limitations on tree size at the global scale[54].

TRAIT CONSTELLATIONS AT THE GLOBAL SCALE. Although exploration of trait PC axes sheds light on the dominant physiological trade-offs structuring tree traits, these first two trait axes account for less than half of overall trait variation (Fig. 2a). To better explore the multidimensional nature of these trade-offs, we subsequently identified groups of traits that form tightly coupled clusters and which reflect distinct aspects of tree function.

Our results show that these 18 traits can be grouped into seven trait constellations, each of which reflects a unique aspect of tree growth, physiology, or ecology (Fig. 4). The largest trait constellation (Fig. 4, pink cluster) loads most heavily on the first trait trade-off (Fig. 4b), capturing various aspects of the leaf economic spectrum. Relationships among specific leaf area, leaf V_{cmax} (the maximum rate of carboxylation), leaf thickness, and leaf nutrient concentrations per mass (N, P, K) are well established[4,6,41,55]. The fact that nearly all leaf traits fall in this cluster (with the exception of leaf area and density) supports the inference that leaf economics represent a unique aspect of tree function that is largely independent of plant size [4].

The second-largest trait constellation (Fig. 4, purple cluster) includes tree-size traits which closely align with the second trade-off, along with the addition of leaf area. As with tree height and canopy size, leaf area directly affects a tree's ability to intercept light down through its canopy[42]. Leaf area

thus serves as an intermediary between the two primary trait axes: it is intrinsically correlated with SLA but exhibits relatively weak correlations with per-mass leaf nutrients. Although the largest discrepancies in leaf area are observed between needle-leaf gymnosperms and broadleaf angiosperms, these trends hold within clades as well, with leaf area among gymnosperms likewise correlating positively with both tree height ($\rho = 0.20$) and crown diameter ($\rho = 0.48$). This cluster thus highlights organismal-level coordination of light interception that integrates tree size, architecture, and leaf shape.

Intermediate to these two largest clusters are three constellations each containing two traits: (1) stem conduit diameter and stomatal conductance (Fig. 4, dark green), capturing organismal-level integration of water transport at the cost of increased desiccation and cavitation risks; (2) stem diameter and bark thickness (Fig. 4, light green), primarily demonstrating intrinsic size-based relationships between stem parts[56,57], which are secondarily related to aridity and fire frequency in some environments[57]; and (3) wood density and leaf density (Fig. 4, orange), indicative of slow/fast life-history strategies, where denser plant parts reduce growth rate and water transport[5,17] but protect against pest damage, desiccation, and mechanical breakage[5,27,39]. Collectively, these two-trait clusters each demonstrate unique and complementary trade-offs that insulate trees against various disturbances and extreme weather events, but at the cost of reduced growth, competitive ability, and productivity under optimal conditions (see Supplemental Discussion).

Lastly, two traits each comprise their own unique cluster: root depth and seed dry mass (Fig. 3, yellow and blue). Root growth is subject to a range of belowground processes (e.g., root herbivory, depth to bedrock) that can promote a disconnect between aboveground climate conditions and belowground traits[54,58,59]. Root depth accordingly has a relatively weak phylogenetic signal ($\lambda = 0.44$) but a strong environmental signal (Figs. 4, S1-S2), reflecting distinct belowground constraints on trait expression. In contrast, seed dry mass exhibits the strongest phylogenetic signal ($\lambda = 0.98$, Fig. 4c) and weakest environmental signal of any trait (Figs. S1-S2). Reproductive traits are subject to unique evolutionary pressures[60], indicative of different seed dispersal vectors (wind, water, animals) and various ecological stressors that uniquely affect seed viability and germination[60]. The emergence of root depth and seed mass as solo functional clusters thus supports previous inference that belowground traits and reproductive traits reflect distinct aspects of tree form and function not captured by leaf or wood trait spectrums.

Collectively, these trait constellations shed light on organismal-level trait coordination and broad-scale differences in trait expression across species and clades. A key challenge in identifying global patterns in trait trade-offs is the relatively sparse trait coverage at the individual level, with only a handful of traits typically measured on any single tree. This limitation is partly due to the enormous range of putative traits that can be measured on trees[6,25]. Here, by using phylogenetic and environmental information to estimate trait expression at the individual level, our approach helps to overcome some of these limitations, enabling us to explore organismal level trade-offs across thousands of species. The benefit of this approach we can include vastly broader phylogenetic, geographic, and trait coverage than would otherwise be possible. Future work, however, should focus on improving the precision of these global trait frameworks by measuring complete sets of traits on individual trees.

To help address these challenges, the seven described trait constellations (Fig. 4) can be used as a starting point for research into organismal-level trait expression in trees. Although the exact subset of traits used in a given study should depend on the intended scope and application, we advise selecting traits which exhibit strong phylogenetic signal and/or low environmental-mediated variation, and ideally have low cross-correlations with other traits in other clusters. These criteria help to ensure that the selected traits can be robustly measured and that they reflect well-defined ecological processes. In line with this, we suggest a baseline set of seven traits selected from these trait constellations: bark thickness, maximum tree height, root depth, specific leaf area, stem conduit diameter, wood density, and seed dry mass. These seven traits represent complementary ecological and evolutionary processes, capturing differences in competitive ability, growth rate, abiotic stress tolerance, reproduction, wood and leaf properties, and above- vs. belowground allocation. Moreover, these traits are relatively well represented in many trait databases and have well-defined definitions and measurement protocols[25,61,62], thus forming a baseline set of reference traits for expanding our understanding of tree functional trait expression.

CONCLUSIONS

Collectively, our analysis reveals key trade-offs and trait constellations governing tree form and function worldwide. We show that tree functional traits predominantly reflect two major functional

trade-offs: one representing leaf-level nutrient-use and photosynthesis, and the other representing competition for light via tree and crown size. Mirroring patterns seen across the entire plant kingdom, these trade-offs capture an ecological gradient from conservative growth strategies under suboptimal environments (cold, dry, frequent disturbances), to acquisitive strategies associated with light competition in high-resource environments (consistently high temperature and water availability). By incorporating traits unique to large woody species, we further identify a unique set of functional constellations and representative traits that reflect the breadth of tree form and function. In doing so, these results elucidate key constraints on functional trait relationships in trees, contributing to our fundamental understanding of the controls on the function, distribution, and composition of forest communities. By identifying a core set of traits that reflect the broad variety of ecological life-history strategies in trees, this work can inform future trait-based research into the functional biogeography of the global forest system.

241 ***Data and code availability***

242 The data and code for replicating the central findings will be made available in a dedicated GitHub
243 repository upon publication.

244

245 ***Author contributions***

246 DSM conceived of this study and analyzed the data, with assistance from LB, CMZ, CA, JvdH,
247 HM, LM, GRS and TWC. Data were contributed by IA, EB, CCFB, JC, BELC, ASD, AG-M, PH,
248 CHL, ASK, ÜN, VDP, JAR, FMS, SS, ACdS, ÊS, PMvB, EW, GB, and JK, who also provided
249 suggestions and feedback on the analyses and interpretations. All authors contributed to the writing
250 and revising of the manuscript.

251

252 ***Competing Interest Statement***

253 The authors declare no competing interests.

254

255 ***Acknowledgements.***

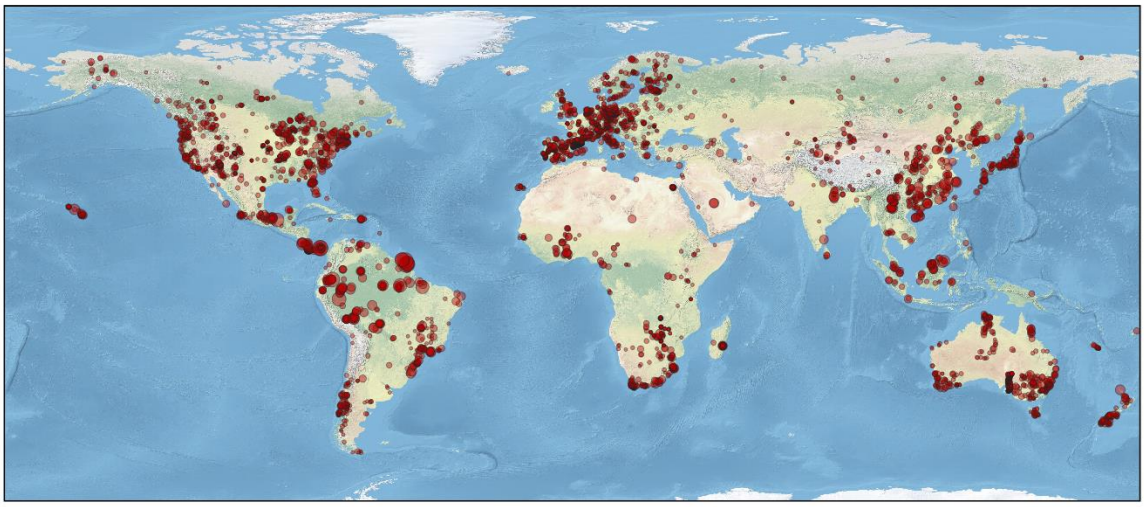
256 This work was primarily funded by the Swiss National Science Foundation, Ambizione grant
257 #PZ00P3_193612 to DS Maynard. Additional support was provided by D.O.B. Ecology to TW
258 Crowther; SNSF Ambizione grant #PZ00P3_193646 to CM Zohner; SNSF Ambizione grant #
259 PZ00P3_179900 to C Averill; U.S. National Science Foundation Graduate Research Fellowship
260 and Embassy of Switzerland in the USA ThinkSwiss Research Fellowship to GR Smith; and
261 UNAM-PAPIIT grant #IN210220 to JA Rosell. The authors thank Owen Atkin and Yongfu Chai
262 for additional data, as well as the TRY database managers and numerous researchers who
263 contributed open-access data.

264

265

266

(a)



(b)

Trait	# Observations	# Species	# Genera	# Locations	Mean	Median	Std. Dev.	Range
Bark thickness (mm)	6209	976	503	172	4.8	3.0	5.1	(0.05 - 53)
Crown height, maximum (m)	5073	266	159	113	5.6	4.7	4.0	(1.0 - 29.9)
Crown diameter, maximum (m)	4136	160	105	98	3.3	2.4	2.7	(1.0 - 24.8)
Leaf area (cm^2)	13362	570	308	440	81.5	42.6	122.9	(0.025 - 999.6)
Leaf density ($g\ cm^{-3}$)	21156	1471	583	256	0.37	0.37	0.12	(0.0025 - 1.3)
Leaf potassium (K) per mass ($mg\ g^{-1}$)	5796	1291	567	269	7.6	6.1	4.7	(0.14 - 35.1)
Leaf nitrogen (N) per mass ($mg\ g^{-1}$)	53837	4588	1275	3505	19.4	18.6	7.5	(2.8 - 51.9)
Leaf phosphorous (P) per mass ($mg\ g^{-1}$)	21812	2819	954	2274	1.2	1.0	0.78	(0.032 - 6)
Leaf thickness (mm)	54228	1821	675	399	0.25	0.23	0.11	(0.005 - 1.9)
Root depth, maximum (m)	1610	363	187	523	2.5	1.5	3.1	(0.25 - 22.2)
Seed dry mass (mg)	564	449	272	73	488.2	27.5	2178.6	(0.094 - 35714)
Specific leaf area ($mm^2\ mg^{-1}$)	64583	4772	1310	2148	16.4	13.5	10.8	(0.56 - 55)
Stem conduit diameter (μm)	491	248	153	106	26.5	22.6	17.3	(4.6 - 110)
Stem diameter, maximum (m)	53591	2044	639	569	0.22	0.17	0.18	(0.1 - 6.0)
Stomatal conductance ($mmol\ m^{-2}\ s^{-1}$)	28465	911	448	248	152.2	121.7	146.0	(1.2 - 1499.7)
Tree height, maximum (m)	36028	1831	639	575	18.8	18.0	8.2	(5.0 - 112.9)
V_{cmax} per mass, at 25°C ($\mu mol\ g^{-1}\ s^{-1}$)	1872	548	313	126	0.47	0.41	0.28	(0.032 - 1.7)
Wood density ($g\ cm^{-3}$)	13713	1655	638	4631	0.58	0.58	0.18	(0.061 - 1.2)

Figure 1. Overview of the 18 functional traits. (a) The unique geographic locations (n = 8490) where tree functional traits were recorded. The size of the circles denotes the relative number of traits measured at each location. (b) The number of unique measurements, species, locations, and genera for each of the 18 traits considered here, along with summary statistics. The analysis included 386,526 trait measurements, encompassing 6905 unique tree species and 1691 unique genera (see Table S1 for corresponding metadata).

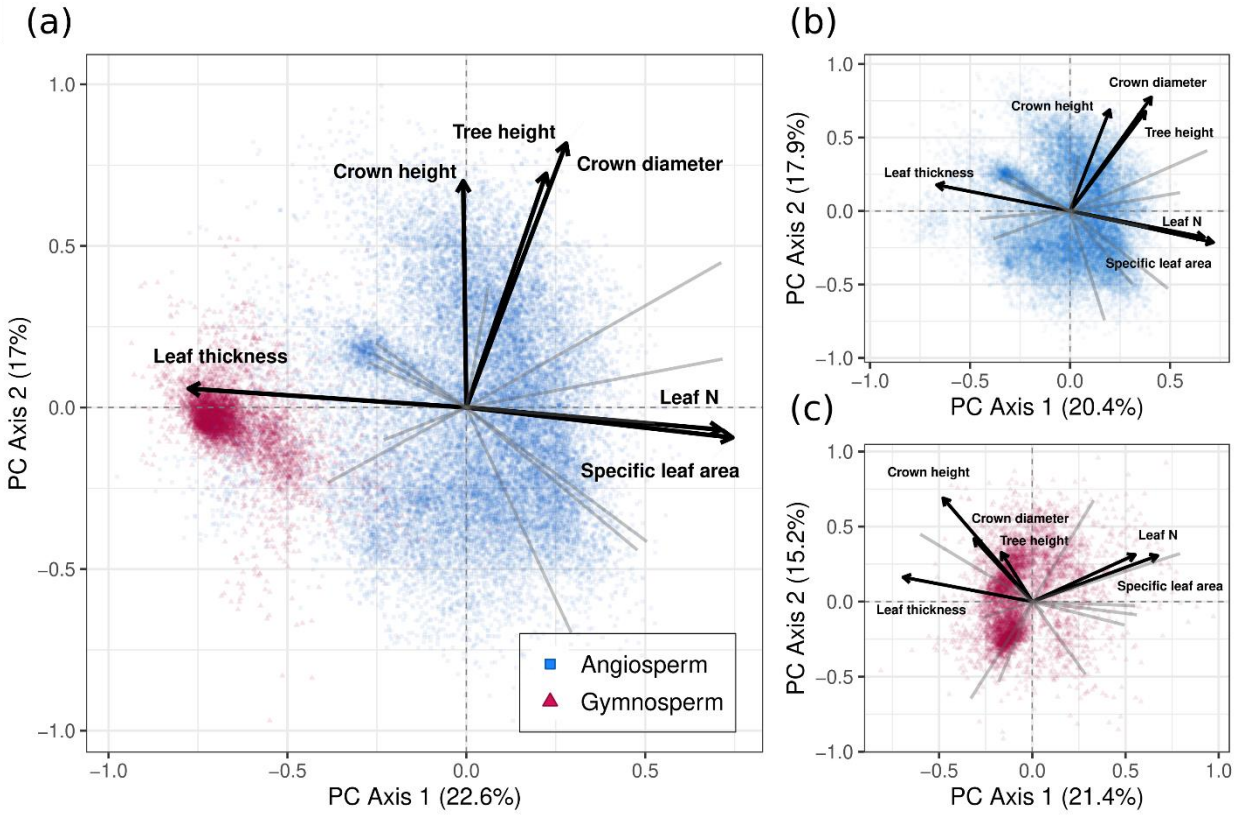


Figure 2. Dominant trait axes and trade-offs. Shown are the first two principal component axes capturing trait trade-offs across the 18 functional traits. (a) All tree species (n = 29,450 observations), (b) angiosperms only (n = 5457), and (c) gymnosperms only (n = 23,993). In (a) the three variables that load most strongly on each axis are show in dark black lines, with the remaining variables shown in light gray. These same six variables are shown in (b) and (c) illustrating how the same trade-offs extend to angiosperms and gymnosperms. See Supplemental Figs. S4-S6 for the full trait PCAs.

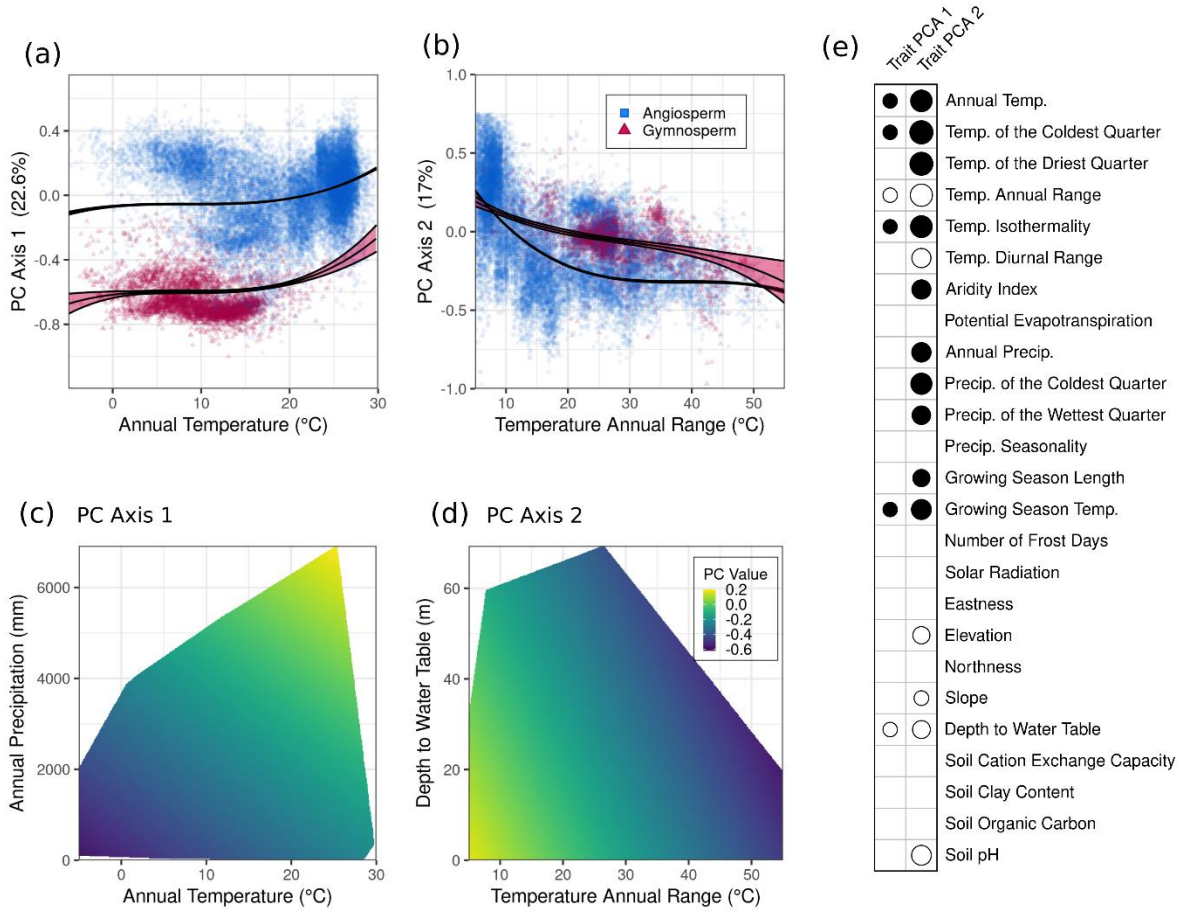


Figure 3. The relationship between environmental variables and trait trade-offs. (a-b) The strongest univariate predictor of each trait trade-off. (c-d) The 2-dimensional surfaces showing the strongest bivariate predictors of each trait trade-off. The surfaces are constrained to the convex hull of the observed variable combinations in the dataset. Note that the x-axes in *c-d* align with those above in *a-b*. (e) Correlations between the first two trait trade-offs and environmental variables. The size of the circle denotes the relative strength of the correlation, with solid circles denoting positive correlations and open circles denoting negative correlations. For clarity, only trait correlations with $|\rho| > 0.2$ are shown. See Fig. S7 for the full set of correlations.

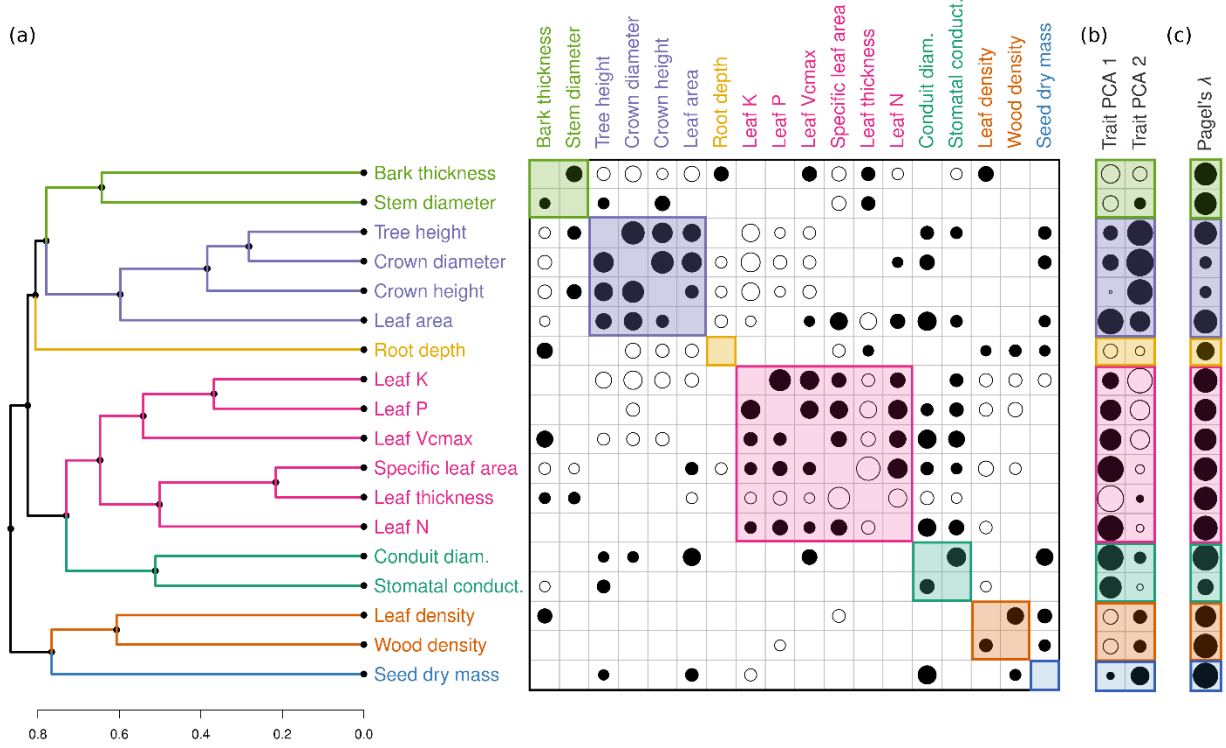


Figure 4. Trait correlations and functional constellations. (a) Trait clusters with high average intra-group correlation. The upper triangle shows species-weighted correlations incorporating intraspecific variation, and the lower triangle gives the corresponding correlations among species-level phylogenetic independent contrasts. The size of the circle denotes the relative strength of the correlation, with solid circles denoting positive correlations and open circles denoting negative correlations. For clarity, only trait correlations with $|\rho| > 0.2$ are shown. (b) Correlations between each trait and each of the first two principal component axes, illustrating which functional trait clusters align most strongly with the dominant axes of trait variation. (c) The species-level phylogenetic signal of each trait, calculated using only the empirical trait measurements.

References

1. Kraft, N. J. B., Godoy, O. & Levine, J. M. Plant functional traits and the multidimensional nature of species coexistence. *Proc. Natl. Acad. Sci.* **112**, 797–802 (2015).

2. Westoby, M. & Wright, I. J. Land-plant ecology on the basis of functional traits. *Trends Ecol. Evol.* **21**, 261–268 (2006).

3. Schmitz, O. J., Buchkowski, R. W., Burghardt, K. T. & Donihue, C. M. *Functional Traits and Trait-Mediated Interactions. Connecting Community-Level Interactions with Ecosystem Functioning. Advances in Ecological Research* vol. 52 (Elsevier Ltd., 2015).

4. Díaz, S. *et al.* The global spectrum of plant form and function. *Nature* **529**, 1–17 (2015).

5. Chave, J. *et al.* Towards a worldwide wood economics spectrum. *Ecol. Lett.* **12**, 351–366 (2009).

6. Maracahipes, L. *et al.* How to live in contrasting habitats? Acquisitive and conservative strategies emerge at inter- and intraspecific levels in savanna and forest woody plants. *Perspect. Plant Ecol. Evol. Syst.* **34**, 17–25 (2018).

7. Rueda, M., Godoy, O. & Hawkins, B. A. Trait syndromes among North American trees are evolutionarily conserved and show adaptive value over broad geographic scales. *Ecography (Cop.).* **41**, 540–550 (2018).

8. Kraft, N. J. B. & Ackerly, D. D. Functional trait and phylogenetic tests of community assembly across spatial scales in an Amazonian forest. *Ecol. Monogr.* **80**, 401–422 (2010).

9. Treurnicht, M. *et al.* Functional traits explain the Hutchinsonian niches of plant species. *Glob. Ecol. Biogeogr.* **29**, 534–545 (2020).

10. Cornwell, W. K. *et al.* Plant traits and wood fates across the globe: rotted, burned, or consumed? *Glob. Chang. Biol.* **15**, 2431–2449 (2009).

11. Greenwood, S. *et al.* Tree mortality across biomes is promoted by drought intensity, lower wood density and higher specific leaf area. *Ecol. Lett.* **20**, 539–553 (2017).

12. van Mantgem, P. J. *et al.* Widespread increase of tree mortality rates in the western United

States. *Science (80-.).* **323**, 521–4 (2009).

13. Butchart, S. H. M. *et al.* Global biodiversity: indicators of recent declines. *Science (80-.).* **328**, 1164–8 (2010).

14. Stahl, U., Reu, B. & Wirth, C. Predicting species' range limits from functional traits for the tree flora of North America. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 13739–13744 (2014).

15. Iverson, L. R. & Prasad, A. M. Predicting Abundance of 80 Tree Species Following Climate Change in the Eastern United States. *Ecol. Monogr.* **68**, 465 (1998).

16. Wright, I. J. *et al.* The worldwide leaf economics spectrum. *Nature* **428**, 821–7 (2004).

17. Reich, P. B. The world-wide 'fast-slow' plant economics spectrum: A traits manifesto. *J. Ecol.* **102**, 275–301 (2014).

18. Adler, P. B. *et al.* Functional traits explain variation in plant life history strategies. *Proceedings of the National Academy of Sciences of the United States of America* vol. 111 740–745 (2014).

19. Poorter, L. *et al.* The importance of wood traits and hydraulic conductance for the performance and life history strategies of 42 rainforest tree species. *New Phytol.* **185**, 481–492 (2010).

20. Wright, S. J. *et al.* Functional traits and the growth-mortality trade-off in tropical trees. *Ecology* **91**, 3664–3674 (2010).

21. Sakschewski, B. *et al.* Leaf and stem economics spectra drive diversity of functional plant traits in a dynamic global vegetation model. *Glob. Chang. Biol.* **21**, 2711–2725 (2015).

22. Petit, R. J. & Hampe, A. Some evolutionary consequences of being a tree. *Annu. Rev. Ecol. Evol. Syst.* **37**, 187–214 (2006).

23. Niinemets, Ü. Global-scale climatic controls of leaf dry mass per area, density, and thickness in trees and shrubs. *Ecology* **82**, 453–469 (2001).

24. 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).

25. Kattge, J. *et al.* TRY plant trait database – enhanced coverage and open access. *Glob.*

Chang. Biol. **26**, 119–188 (2020).

26. Givnish, T. J. Plant stems: biomechanical adaptation for energy capture and influence on species distributions. *Plant stems Physiol. Funct. Morphol.* 3–49 (1995).

27. King, D. A., Davies, S. J., Tan, S. & Noor, N. S. M. The role of wood density and stem support costs in the growth and mortality of tropical trees. *J. Ecol.* **94**, 670–680 (2006).

28. Karger, D. N. *et al.* Climatologies at high resolution for the earth's land surface areas. *Sci. Data* **4**, 1–20 (2017).

29. Karger, D. N., Conrad, O., Böhner, J., Kawohl, T. & Kreft, H. Data from: Climatologies at high resolution for the earth's land surface areas. *Dryad Digit. Repos.* (2018).

30. Trabucco, A. & Zomer, R. Global Aridity Index and Potential Evapotranspiration (ET0) Climate Database v2. *Figshare. https://doi.org/10.6084/m9.figshare.7504448.v3* (2019).

31. Obu, J. *et al.* Northern Hemisphere permafrost map based on TTOP modelling for 2000–2016 at 1 km2 scale. *Earth-Science Rev.* **193**, 299–316 (2019).

32. Hengl, T. *et al. SoilGrids250m: Global gridded soil information based on machine learning*. *PLoS ONE* vol. 12 (2017).

33. Tuanmu, M. N. & Jetz, W. A global, remote sensing-based characterization of terrestrial habitat heterogeneity for biodiversity and ecosystem modelling. *Glob. Ecol. Biogeogr.* **24**, 1329–1339 (2015).

34. Fan, Y., Li, H. & Miguez-Macho, G. Global patterns of groundwater table depth. *Science (80-.).* **339**, 940–943 (2013).

35. Diniz-Filho, J. A. F., De SanT'Ana, C. E. R. & Bini, L. M. An eigenvector method for estimating phylogenetic inertia. *Evolution (N. Y).* **52**, 1247–1262 (1998).

36. Swenson, N. G. Phylogenetic imputation of plant functional trait databases. *Ecography (Cop.).* **37**, 105–110 (2014).

37. Poorter, H., Niinemets, Ü., Poorter, L., Wright, I. J. & Villar, R. Causes and consequences of variation in leaf mass per area (LMA): A meta-analysis. *New Phytol.* **182**, 565–588 (2009).

385 38. Firn, J. *et al.* Leaf nutrients, not specific leaf area, are consistent indicators of elevated
386 nutrient inputs. *Nat. Ecol. Evol.* **3**, 400–406 (2019).

387 39. Onoda, Y. *et al.* Global patterns of leaf mechanical properties. *Ecology Letters* vol. 14
388 301–312 (2011).

389 40. Tian, D. *et al.* Global leaf nitrogen and phosphorus stoichiometry and their scaling
390 exponent. *Natl. Sci. Rev.* **5**, 728–739 (2018).

391 41. Walker, A. P. *et al.* The relationship of leaf photosynthetic traits - Vcmax and Jmax - to
392 leaf nitrogen, leaf phosphorus, and specific leaf area: A meta-analysis and modeling study.
393 *Ecol. Evol.* **4**, 3218–3235 (2014).

394 42. Niinemets, Ü. A review of light interception in plant stands from leaf to canopy in
395 different plant functional types and in species with varying shade tolerance. *Ecol. Res.* **25**,
396 693–714 (2010).

397 43. Harja, D., Vincent, G., Mulia, R. & van Noordwijk, M. Tree shape plasticity in relation to
398 crown exposure. *Trees - Struct. Funct.* **26**, 1275–1285 (2012).

399 44. Poorter, L., Bongers, L. & Bongers, F. Architecture of 54 moist-forest tree species: Traits,
400 trade-offs, and functional groups. *Ecology* **87**, 1289–1301 (2006).

401 45. Simard, M., Pinto, N., Fisher, J. B. & Baccini, A. Mapping forest canopy height globally
402 with spaceborne lidar. *J. Geophys. Res. Biogeosciences* **116**, 1–12 (2011).

403 46. Klein, T., Randin, C. & Körner, C. Water availability predicts forest canopy height at the
404 global scale. *Ecol. Lett.* **18**, 1311–1320 (2015).

405 47. Sperry, J. S., Hacke, U. G. & Pittermann, J. Size and function in conifer tracheids and
406 angiosperm vessels. *Am. J. Bot.* **93**, 1490–1500 (2006).

407 48. Choat, B. *et al.* Global convergence in the vulnerability of forests to drought. *Nature* **491**,
408 752–755 (2012).

409 49. Moles, A. T. *et al.* Which is a better predictor of plant traits: Temperature or precipitation?
410 *J. Veg. Sci.* **25**, 1167–1180 (2014).

411 50. Mittelbach, G. G. *et al.* Evolution and the latitudinal diversity gradient: Speciation,
412 extinction and biogeography. *Ecol. Lett.* **10**, 315–331 (2007).

51. Schemske, D. W., Mittelbach, G. G., Cornell, H. V., Sobel, J. M. & Roy, K. Is There a Latitudinal Gradient in the Importance of Biotic Interactions? *Annu. Rev. Ecol. Evol. Syst.* **40**, 245–269 (2009).

52. Olson, M. E. *et al.* Plant height and hydraulic vulnerability to drought and cold. *Proc. Natl. Acad. Sci. U. S. A.* **115**, 7551–7556 (2018).

53. Zanne, A. E. *et al.* Functional biogeography of angiosperms: life at the extremes. *New Phytol.* **218**, 1697–1709 (2018).

54. Fan, Y., Miguez-Macho, G., Jobbágy, E. G., Jackson, R. B. & Otero-Casal, C. Hydrologic regulation of plant rooting depth. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 10572–10577 (2017).

55. Tripler, C. E., Kaushal, S. S., Likens, G. E. & Todd Walter, M. Patterns in potassium dynamics in forest ecosystems. *Ecol. Lett.* **9**, 451–466 (2006).

56. Poorter, L., Mcneil, A., Hurtado, V. H., Prins, H. H. T. & Putz, F. E. Bark traits and life-history strategies of tropical dry- and moist forest trees. *Funct. Ecol.* **28**, 232–242 (2014).

57. Rosell, J. A. Bark thickness across the angiosperms: More than just fire. *New Phytol.* **211**, 90–102 (2016).

58. Weemstra, M. *et al.* Towards a multidimensional root trait framework: a tree root review. *New Phytol.* **211**, 1159–1169 (2016).

59. Laughlin, D. C. *et al.* Root traits explain plant species distributions along climatic gradients yet challenge the nature of ecological trade-offs. *Nat. Ecol. Evol.* (2021) doi:10.1038/s41559-021-01471-7.

60. Moles, A. T. *et al.* Factors that shape seed mass evolution. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 10540–10544 (2005).

61. Pérez-Harguindeguy, N. & Díaz, S. New handbook for standardised measurement of plant functional traits worldwide. *Aust. J. ...* **61**, 167–234 (2013).

62. Cornelissen, J. H. C. *et al.* A handbook of protocols for standardised and easy measurement of plant functional traits worldwide. *Aust. J. Bot.* **51**, 335 (2003).

Materials & Methods

Trait information

Trait data were obtained from the TRY plant trait database[1] in April 2020. Data were cleaned by converting all traits to standardized units and by matching species names to The Plant List (TPL) database v1.1 (http://www.theplantlist.org, accessed June 2020) using the *Taxonstand* package in R v3.6.0[2]. Synonyms were replaced with accepted names, when available. The phylogenetic tree was taken from the seed plant phylogeny of Smith & Brown (2018), and species names were likewise cleaned and harmonized using the TPL database. To limit our analysis to tree functional traits only, we used the BGCI GlobalTreeSearch database v1.3[4,5], containing a comprehensive list of ca. 60,000 tree species compiled and harmonized from across 500 sources. The BGCI database uses TPL for much of its taxonomic identification, but to ensure consistency among all sources we used the same name harmonization pipeline as with TRY and the seed plant phylogeny. We constrained the set of traits and the phylogenetic tree to those species that could be matched to the BGCI database (n = 54,153 species matched), and we likewise trimmed the phylogenetic tree to the set of species matched in TRY. We further excluded any trait observations that did not have corresponding geographic coordinates.

Traits were selected based on data availability, phylogenetic and spatial coverage, and importance for tree growth, survival, and competition. We prioritized traits that are commonly used in other leaf and wood economic spectrums[6], thus focusing on leaf traits measured per unit mass rather than unit area, where possible. Secondly, we prioritized traits which are important indicators of tree growth and structure, such as stem conduit diameter and crown dimensions. Thirdly, we selected traits which capture important ecological processes and life-history differences among trees (e.g., root depth). This process resulted in 30 putative traits for analysis (Table S1, Supplemental Data References). Of these, we omitted traits with low sample sizes (either overall, or at the species, genus, or geographic level), those that are redundant or intrinsically correlated with other traits, and those that are known to be highly sensitive to trait assay conditions but were not clearly standardized. This pruning resulted in a set of 18 focal traits for use in the final analysis, with the 12 omitted traits used only to improve predicted power via trait covariation (see model

process, below). Trait values were converted to common units where necessary (e.g., mm to cm). For each trait, we selected sub-categories (as given by TRY) that denoted comparable measurements and reflected uniform assay conditions (e.g., V_{cmax} measured at 25°C) (Table S1).

Model details

In order to consider trait trade-offs at the organismal level which accounted for intraspecific variation, we used machine learning models to estimate all 18 trait values for each individual tree in each location (Fig. 1a). We modeled trait expression as a function of both environmental and phylogenetic information so as to estimate traits with weak phylogenetic signals but strong abiotic filtering, and to incorporate intraspecific variation and ontogenetic plasticity into our analysis. Specifically, we used random forest (RF) models to estimate trait values for each observation using the *ranger* package in R[7]. Initial exploration showed negligible effects of model tuning, such that the default hyper-parameters were used to prevent overfitting. In order to minimize the influence of data-recording errors or unit mismatches in the dataset, trait values which occurred outside of the bulk of the trait distribution were investigated as outliers. Those which could not be externally verified and which were biologically unreasonable were removed (e.g., stem diameters >15 m). When modeling tree height, canopy size, and root depth, we only considered observations with height >5 m, diameter >10 cm, root depth >25 cm, and canopy dimensions >1 m high and wide[8], thereby ensuring that our analysis focused on adult trees rather than saplings or woody shrubs. We subsequently implemented quantile random forest[9,10] to estimate the upper 90th percentile trait value for maximum stem diameter, canopy dimensions, and root depth. In all other cases the imputed traits represent the mean predicted value across the random forest.

Environmental covariates used in the models included 50 variables encompassing range of climate[11–13], soil[14], topographic[15], and geological[16] variables (Table S2). We omitted variables that directly measure plant community composition or biotic factors (e.g., NDVI or % forest cover) so as to ensure the resulting geographic layers solely encompassed abiotic factors. Layers were sampled from a previously prepared global composite (see van den Hoogen et al. 2019 for details). Briefly, all covariate map layers were resampled and reprojected to a unified pixel grid in EPSG:4326 (WGS84) at 30 arcsec resolution (approximately 1 km^2 at the equator). Layers with a higher original pixel resolution were downsampled using a mean aggregation method; layers with a lower original resolution were resampled using simple upsampling (that is, without interpolation)

to align with the higher resolution grid. The set of environmental covariates for each trait measurement was obtained by sampling this composite image at each unique latitude and longitude value given in the TRY database.

Phylogenetic information was incorporated in the form of phylogenetic eigenvectors[18–21]. We first calculated the pairwise cophenetic phylogenetic distance matrix across all 54,153 tree species that could be matched to both the BGCI tree list and the plant phylogeny. This matrix was then double-centered by rows and columns[21,22], and the first 50 orthogonal eigenvectors were extracted from this matrix for use as continuous predictors in the random forest models. The choice of 50 eigenvectors (out of 54,153 possible) was in line with previous analyses to prevent over-fitting and to ensure the model was identifiable[22]. This also resulted in the same number of environmental and phylogenetic predictor variables.

To leverage trait covariation among the disparate observations, we used a two-step algorithm to improve predictive power and imputation accuracy[20,23]. First, following standard approaches[6,24–26], trait values were log-transformed, allowing for comparisons across trait distributions which are highly right-skewed and vary by several orders of magnitude[26] (Fig. 1). Using the general approach of Stekhoven & Bühlmann (2012), we next implemented a random forest on all traits for all observations. We then used these initial models to predict the full set of trait values for each observation (including the 12 ancillary traits not included in the focal analysis, Table S1). We then refit the random forest models for each trait, using the full set of predicted traits (apart from the focal traits) as covariates. For the final analysis, observed traits were used in place of imputed traits, when available, with the exception of maximum tree height, stem diameter, root depth, and crown size, where the upper 90th percentile trait values were used. Variable importance in the random forest models was calculated using the "permutation" metric, reflecting the variance in responses across predictors[7].

Model performance

Model performance was quantified using buffered leave-one-out cross-validation[27]. To avoid overfitting, we followed the approach recommended in Roberts *et al.* (2017) and fit a simple linear model to the data, where trait expression was modeled as a linear function of phylogenetic and environmental covariates. We then assessed spatial autocorrelation of the residuals using Moran's

I plots using the *ncf* package in R, which displays the value of spatial autocorrelation (ranging from -1 to 1) as a function of distance[28]. We likewise assessed residual phylogenetic autocorrelation across taxonomic ranks (genus, family, order, group), using the the *ape* package in R. In general, spatial autocorrelation was low (I<0.10) (Fig. S8), with the exception of leaf phosphorous, which exhibited slight autocorrelation up to ~250 km. Residual phylogenetic correlation was likewise low, and generally only observable at the genus level, apart from crown size and conduit diameter, which exhibited residual autocorrelation up to the family level (Fig. S9). Thus, to be conservative, for all traits except crown size and conduit diameter, we used a genus-level spatial buffer of 250 km to exclude test/training data; and for crown size and conduit diameter we used a family-level buffer at 250 km. To implement the cross-validation accuracy assessment, we first randomly selected a focal species, with the out-of-fit test data containing all observations for that species for the focal trait. To construct the corresponding training data, we excluded all observation of the same genus (or family) that fell within a 250km spatial buffer of any of the training points for that species. The random forest models were then fit using the buffered training data, and used to predict the average trait value for the omitted species[27]. This procedure was repeated for each unique species for each trait, up to 1000 times, with a randomly sampled focal species selected at each iteration.

Predictive accuracy was assessed in two ways. First, following the recommendation of Li (2017), we calculated the cross-validated coefficient of determination relative to the 1:1 line (termed "VEcv", Li 2017), which provides a normalized version of the mean-squared-error (MSE) that allows for comparisons across data types and units. Specifically, this value is calculated as: R^2_{VEcv} = 1 - $\sum(y_i^{pred} - y_i^{obs})^2 / \sum(y_i^{obs} - \bar{y})^2$ = 1 – SSE/TSS = 1 – MSE / $\hat{\sigma}^2$, where the summation is taken across the species, and the predicted values are estimated out-of-fit using the buffered cross-validation procedure outlined above. Importantly, this metric is not the same as a regression-based goodness-of-fit, as it is calculated by direct comparison of observed vs. predicted values[29]. Second, we also report the median absolute percentage error (MdAPE), which gives a more interpretable estimate of the expected error a given prediction, calculated as MdAPE = $median\left(\left| y_i^{obs} - y_i^{pred} \right| / y_i^{obs} \right) \times 100$. Although the models were fit using log-transformed data, accuracy was assessed on the non-logged values in their original units.

Principal component analysis

Species-weighted principal component analysis (PCA) was conducted on the full set of imputed traits using the *aroma.light* package in R. The weights were set to be inversely proportional to the number of observations for each species, which allowed us to incorporate intraspecific variation while also ensuring that each species had the same overall contribution to global trade-offs. Representative vectors for each axis were identified by selecting those that correlated most uniquely on each of the first two principle component axes.

Abiotic relationships

To identify univariate and bivariate relationships among trait trade-offs and environmental conditions, we first identified the environmental variable that correlated most strongly with each of the two PC axes (Fig. 3a-b, Fig. S7). We used Spearman rank correlations to allow for nonlinear relationships among traits and environmental conditions. To visualize these correlations, we separately fit third-order monotonic regression polynomials for angiosperms and gymnosperms, and obtained 95% bootstrap confidence intervals by randomly sampling one observation for each species per iteration, repeated 500 times. To explore the bivariate predictors of trait trade-offs, we then fit a series of simple pairwise linear regression models to identify which additional environmental variable led to the highest subsequent increase in explanatory power for each trait (measured via adjusted R^2). To avoid spurious relationships due to the large number of pairwise combinations (50 choose 2), we only considered a subset of representative environmental traits, identified via cluster analysis: annual precipitation, annual temperature, temperature annual range, precipitation seasonality, aridity index, growing season mean temperature, growing season length, permafrost extent, soil water-holding capacity, soil cation-exchange capacity, soil pH, topographic northness, topographic eastness, and depth to water table. To visualize the resulting patterns (Fig. 3c-d), we plotted the smooth regression surfaces across the full range of environmental conditions, restricted to the convex hull of the observed variable combinations in the dataset.

Hierarchical cluster analysis

Trait cluster analysis was conducted using hierarchical clustering on the species-level correlation matrix. First, we calculated species-weighted rank correlations between pairs of traits using the *wCorr* package in R, which again allowed us to incorporate intraspecific trait variation while

ensuring each species contributed equal weight. The optimal number of clusters was identified using the silhouette method in the *dendextend* package in R, and the dendrogram was subsequently cut into clusters based on groups of traits which exhibited consistently high average intra-group correlation. As an alternate measure of trait correlation which accounts for phylogenetic relatedness, we calculated phylogenetic independent contrasts[30] on species-level average trait values using the *ape* package. The corresponding correlations among these contrasts are shown in the bottom triangle of the correlation matrix in Fig. 4. Species-level phylogenetic conservatism was calculated via Pagel's λ, using only the empirically measured values in the TRY dataset.

All analyses were conducted in R v. 3.6.0, with the exception the phylogenetic eigenvector calculations, which were obtained using the *Arpack* package in Julia v. 1.6.2.

Methods References

1. Kattge, J. *et al.* TRY plant trait database – enhanced coverage and open access. *Glob. Chang. Biol.* **26**, 119–188 (2020).

2. R Core Team. *R: A Language and Environment for Statistical Computing*. (R Foundation for Statistical Computing, 2021).

3. Smith, S. A. & Brown, J. W. Constructing a broadly inclusive seed plant phylogeny. *Am. J. Bot.* **105**, 302–314 (2018).

4. BGCI. *GlobalTreeSearch online database (version 1.4). Botanic Gardens Conservation International*. (Available at https://tools.bgci.org/global_tree_search.php, 2020).

5. Beech, E., Rivers, M., Oldfield, S. & Smith, P. P. GlobalTreeSearch: The first complete global database of tree species and country distributions. *J. Sustain. For.* **36**, 454–489 (2017).

6. Díaz, S. *et al.* The global spectrum of plant form and function. *Nature* **529**, 1–17 (2015).

7. Wright, M. N. & Ziegler, A. Ranger: A fast implementation of random forests for high dimensional data in C++ and R. *J. Stat. Softw.* **77**, 1–17 (2017).

8. Gschwantner, T. *et al.* Common tree definitions for national forest inventories in Europe. *Silva Fenn.* **43**, 303–321 (2009).

9. Meinshausen, N. Quantile Regression Forests. *J. Mach. Learn. Res.* **7**, 983–999 (2006).

10. Fox, E. W., Ver Hoef, J. M. & Olsen, A. R. Comparing spatial regression to random forests for large environmental data sets. *PLoS One* **15**, 1–22 (2020).

11. Karger, D. N. *et al.* Climatologies at high resolution for the earth's land surface areas. *Sci. Data* **4**, 1–20 (2017).

12. Karger, D. N., Conrad, O., Böhner, J., Kawohl, T. & Kreft, H. Data from: Climatologies at high resolution for the earth's land surface areas. *Dryad Digit. Repos.* (2018).

13. Trabucco, A. & Zomer, R. Global Aridity Index and Potential Evapotranspiration (ET0) Climate Database v2. *Figshare. https://doi.org/10.6084/m9.figshare.7504448.v3* (2019).

14. Hengl, T. *et al. SoilGrids250m: Global gridded soil information based on machine learning*. *PLoS ONE* vol. 12 (2017).

15. Tuanmu, M. N. & Jetz, W. A global, remote sensing-based characterization of terrestrial habitat heterogeneity for biodiversity and ecosystem modelling. *Glob. Ecol. Biogeogr.* **24**, 1329–1339 (2015).

16. Fan, Y., Li, H. & Miguez-Macho, G. Global patterns of groundwater table depth. *Science (80-.).* **339**, 940–943 (2013).

17. van den Hoogen, J. *et al.* Soil nematode abundance and functional group composition at a

global scale. *Nature* **572**, 194–198 (2019).

18. Swenson, N. G. Phylogenetic imputation of plant functional trait databases. *Ecography (Cop.).* **37**, 105–110 (2014).

19. Bauman, D., Drouet, T., Dray, S. & Vleminckx, J. Disentangling good from bad practices in the selection of spatial or phylogenetic eigenvectors. *Ecography (Cop.).* **41**, 1638–1649 (2018).

20. Penone, C. *et al.* Imputation of missing data in life-history trait datasets: Which approach performs the best? *Methods Ecol. Evol.* **5**, 1–10 (2014).

21. Diniz-Filho, J. A. F., De SanT'Ana, C. E. R. & Bini, L. M. An eigenvector method for estimating phylogenetic inertia. *Evolution (N. Y).* **52**, 1247–1262 (1998).

22. Diniz-Filho, J. A. F., Rangel, T. F., Santos, T. & Mauricio Bini, L. Exploring patterns of interspecific variation in quantitative traits using sequential phylogenetic eigenvector regressions. *Evolution (N. Y).* **66**, 1079–1090 (2012).

23. Stekhoven, D. J. & Bühlmann, P. Missforest-Non-parametric missing value imputation for mixed-type data. *Bioinformatics* **28**, 112–118 (2012).

24. Reich, P. B. The world-wide 'fast-slow' plant economics spectrum: A traits manifesto. *J. Ecol.* **102**, 275–301 (2014).

25. 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).

26. Kattge, J. *et al.* TRY - a global database of plant traits. *Glob. Chang. Biol.* **17**, 2905–2935 (2011).

27. Roberts, D. R. *et al.* Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography (Cop.).* **40**, 913–929 (2017).

28. Dormann, C. F. *et al.* Methods to account for spatial autocorrelation in the analysis of species distributional data: A review. *Ecography (Cop.).* **30**, 609–628 (2007).

29. Li, J. Assessing the accuracy of predictive models for numerical data: Not r nor r2, why not? Then what? *PLoS One* **12**, 1–16 (2017).

30. Felsenstein, J. Phylogenies and the Comparative Method. *Am. Nat.* **125**, 1–15 (1985).