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

Flores Moreno, Habacuc; Fazayeli, Farideh; Banerjee, Arindam; [et al.]. «Robustness of trait connections across environmental gradients and growth forms». *Global ecology and biogeography*, Vol. 28, issue 12 (Dec. 2019), p. 1806-1826. DOI 10.1111/geb.12996

This version is available at <https://ddd.uab.cat/record/218224>

under the terms of the  **CC BY** COPYRIGHT license

Robustness of trait connections between multiple plant organs across environmental gradients and growth forms

*Habacuc Flores-Moreno^{1,2}, *Farideh Fazayeli³, Arindam Banerjee³, Abhirup Datta⁴, Jens Kattge^{5,6}, Ethan E. Butler², Owen K. Atkin^{7,8}, Kirk Wythers², Ming Chen², Madhur Anand⁹, Michael Bahn¹⁰, Sabina Burrascano¹¹, Chaeho Byun¹², J. Hans C. Cornelissen¹³, Joseph Craine¹⁴, Andres Gonzalez-Melo¹⁵, Wesley N. Hatttingh¹⁶, Steven Jansen¹⁷, Nathan J.B. Kraft¹⁸, Koen Kramer¹⁹, Daniel C. Laughlin²⁰, Vanessa Minden²¹, Ülo Niinemets²², Vladimir Onipchenko²³, Josep Peñuelas^{24,25}, Nadejda A. Soudzilovskaia²⁶ and Peter B. Reich^{2,27}

*Authors would like to be considered joint first authors, and for this article to be cited as Flores-Moreno, H and Fazayeli, F, et al.

Author for correspondence: Habacuc Flores-Moreno. Email: habacucfm@gmail.com

Acknowledgments

HFM, PBR, MC, EEB, KRW, AB were funded by United States Department of Energy (DE-SL0012677). HFM was partially funded by a University of Minnesota Institute on the Environment Discovery Grant. AB and PBR were supported by NSF grant IIS-1563950. PBR was funded by two University of Minnesota Institute on the Environment Discovery Grants. O.K.A. acknowledges funding by the Australian Research Council (CE140100008). KK was supported by the EU-FP7 project BACCARA (226299) and the national project Resilient Forests (KB-29-009-003). JP acknowledges the financial support from the European Research Council Synergy grant ERC-SyG-2013-610028 IMBALANCE-P. V.O. thanks Russian Science Foundation (RSF) for financial support (# 14-50-00029). The study has been supported by the TRY initiative on plant traits (<http://www.try-db.org>). The TRY initiative and database is hosted, developed and maintained at the Max Planck Institute for Biogeochemistry, Jena, Germany. TRY is currently supported by DIVERSITAS/Future Earth and the German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena- Leipzig. Discussions with Rhiannon Dalrymple, Santiago Soliveres and Matthew Michalska-Smith contributed greatly to this paper.

Biosketch: Habacuc Flores-Moreno undertook this research as part of a postdoc in Peter Reichs' lab at University of Minnesota. His work focuses on linking the role of plant functional traits to community and ecosystem process in different environmental and ecological contexts.

Figures: 2 (Colour figures: Figure 1 and 2)

Tables: 2

Robustness of trait connections between multiple plant organs across environmental gradients and growth forms

Short running title: Global variation on plant trait connections

Abstract

* **Aim:** Plant trait databases often contain strong trait-trait correlations, where direct and indirect trait connections may be confounded. This confounding of direct connections (undirected statistical dependency) and indirect connections (mediated by other traits) may hinder our understanding of how direct connections between traits relate to and reflect plant strategies. We identified direct connections across plant traits relevant to competition, resource and reproductive strategies and explored whether these connections, within and between organs, vary across climate regions and growth forms.

* **Location:** Global.

* **Major taxa studied:** Plants.

* **Methods:** We used probabilistic graphical models and a database of ten plant traits (specific leaf area, mass- and area-based leaf nitrogen and phosphorous content, leaf lifespan, leaf area, plant height, stem specific density and seed mass) for 16,281 species to describe direct and indirect connections for woody and non-woody plants across tropical, temperate, arid, cold and polar regions.

* **Results:** Land plants have high direct connectivity across all traits and organs; leaf lifespan and stem specific density share connections with all other traits. For both growth forms, two groups of traits form distinct units: one related to resource acquisition, the other to plant architecture and reproduction. Across both growth forms, we observed less trait connectivity in climates with more seasonal variation in precipitation or temperature. Woody species had higher connectivity across traits within organs (higher modularity) in polar compared to temperate and tropical climates, while non-woody species did not show significant differences in modularity across climate regions.

* **Main conclusion:** Plant traits across organs are highly connected. However, connectivity among trait is also shaped by environmental conditions. Differences in modularity patterns across growth forms reflect alternative strategies that plants use to cope with local environmental conditions.

Keywords: global patterns in plant strategies; plant functional traits; plant strategies integration; traits-interdependence; trait networks

Introduction

Plant traits are not independent of each other biologically or statistically. An accurate description of their interdependency gives us a clearer view of a myriad of processes, from the links between morphological traits and physiological function inside the leaves of plants (Osnas et al., 2013; Westoby et al., 2013; Poorter et al., 2014), to ecosystem-scale processes such as global gross primary production (Wang et al., 2012). Evidence of generalities in the coordination of traits exist, but it is mostly limited to single organs or few traits at broad spatial scales (Wright et al., 2004; Chave et al., 2009; Moles et al., 2009), or across organs but only for certain regions and/or growth forms (Ackerly, 2004; Baraloto et al., 2010; Freschet et al., 2010; Liu et al., 2010; Fortunel et al., 2012; Cheng et al., 2015; Kramer-Walter et al., 2016). Diaz et al. (2016) found strong evidence for coordination (i.e. non-random variation between traits) among six traits measured on different organs at global scales. As of yet, however, it remains unknown whether the connections among traits across organs are similar among species subgroups and across environments.

We use ‘connection’ in a specific way in this manuscript; as short for undirected statistical dependencies among traits (Shiple, 2016). A connection represents a direct linkage between two traits arising as a result of either conditional dependence between these traits (sometimes, and perhaps often, functional) or can be the outcome of a set of traits sharing a connection to a third common trait. Meanwhile, missing connections are due to conditional independence among traits where no direct connection between the two traits exists, and any link that may exist between them is indirectly mediated by another trait. We evaluate the connections among traits through a test of undirected statistical independence that accounts for co-variation with all other traits in the data set. This approach thus allows us to address (i) whether correlations among traits all represent direct connections or whether some are mediated by or due to co-variance with other traits, and (ii) whether weakness in simple bivariate correlations might mask connections. We then assess whether and how such connections collectively describe networks across scales of broad climate regions and plant life forms.

Previous studies on multi-organ, multi-trait datasets have typically been limited in geographic scope and described the main axes of trait variation, the association of traits to each of these axes, and associated trait correlations and trade-offs (e.g. Ackerly, 2004; Freschet et al., 2010; Kramer-Walter et al., 2016). Usually, two axes are defined; one has been interpreted as a resource uptake axis, and a second axis as related to competitive ability, and/or response to disturbance (Wright et al., 2007; Baraloto et al., 2010). Evidence suggests that the trait composition of these axes changes across communities under different environmental conditions. For instance, the independence of root, leaf, and stem traits varies across environmental gradients, with plant communities from wetter environments showing the leaf axis to be orthogonal to the root and stem axes (Baraloto et al., 2010; Fortunel et al., 2012; Jager et al., 2015), while in drier environments orthogonality across these organs is lost (Laughlin et al., 2010; Li & Bao, 2015; de la Riva et al., 2016). Similarly, the independence of traits related to

water economy from the resource acquisition axis increases in environments with higher precipitation (Ackerly, 2004; Ishida et al., 2008; Baraloto et al., 2010; Fortunel et al., 2012; de la Riva et al., 2016). Given differences among floras in the number of main axes of trait variation, the trait composition of these axes, and independence among organs, it is still unclear whether the integration (i.e. level of cross-linked trait connection across the plant) of traits changes in a predictable way across broad environmental gradients, or whether multiple strategies for different environmental conditions or plant types exist.

Strong integration across traits across different organs is expected, as matching tissue strategies should be advantageous at the whole-plant scale (Reich, 2014). However, integration across traits among organs may not always be advantageous, as in local environments not all resources necessary for a plant to persist and grow may be available in the same relative supply at the same time (Bloom et al., 1985). Under this scenario, high integration among traits may not always be advantageous for plants. This may be reflected in differences in the connection among traits across environments. We hypothesize that higher seasonal variation will favour less connectivity across traits, as more modular, less highly connected networks may be more stable against external perturbations (Alon, 2003; Gilarranz et al., 2017). On the other hand, woody species have a long-persisting, reinforced stem aboveground that allow vertical and lateral expansion (via cambium inside trunks or broadly branching canopies), while non-woody species lack such a stem. These crucial differences across woody and non-woody strategies has profound impacts on the phenology, reserve patterns, and biophysical requirements of these groups, which may also have implications on the integration among traits of these groups of species. Connectivity among network components represents a compromise between efficiency and cost of connections, where more connected networks are more efficient, but more costly to maintain (Alon, 2003). Thus, we expect that non-woody species may display higher connectivity among traits that sustain their high growth potential and capitalize on favourable conditions, compared to woody species which have lower growth potential, but better tolerate unfavourable conditions (Larcher, 2003; Reich, 2014). Overall, we hypothesize that connections among traits across organs will be higher across non-woody species compared to woody species and will be lower in environments with higher seasonal variation and lower water availability.

An equally important challenge in understanding connections among traits is interpreting evidence from correlative analyses and dimension reduction techniques. Commonly used techniques such as principal component analyses (PCA) effectively decrease the dimensionality of multivariate trait datasets. However, the biological interpretation of the principal component axes (PC axes) and of the traits composing an axis requires a certain level of intuition, especially as the number of axes considered increases (Peres-Neto et al., 2003; Legendre & Legendre, 2012). Further, standard practices such as only selecting a subset of PC axes prevents a full exploration of the multivariate nature of plant strategies and can have unintended consequences (e.g. erroneous inference due to the sole focus on the first few axes of the multivariate space; Uyeda et al., 2015). On the other hand, while correlative analyses generally capture the marginal dependency among variables (Li & Gui, 2006; Legendre & Legendre, 2012), it is impossible to detect indirect from direct dependencies among variables from

this type of analysis alone (Westoby et al., 1995; Shipley et al., 2006; Cai et al., 2012; Ebert-Uphoff & Deng, 2012; Osnas et al., 2013). Distinguishing direct and indirect connections among traits is necessary for understanding the mechanistic roots of trait correlations that define plant strategies, and can thus help clarify the causal link between traits and fitness components (Shipley et al., 2006), connections among traits and function (Poorter et al., 2014), and the role that traits play in influencing higher-level processes and vegetation attributes (e.g. RGR, NPP; Reich et al., 1992).

To address our hypotheses we focus on the following overarching question: how do connections among traits vary across growth forms (woody and non-woody) and environmental gradients? We (i) describe the global trait network (undirected statistical dependency between ten functional traits) among 16,281 plant species from sites around the world; (ii) compare the differences in the trait networks of woody and non-woody species; and (iii) assess how the trait network of woody and non-woody species changes across five broad climate regions (tropical, temperate, arid, cold and polar).

Methods

Trait, taxonomic data and climate zones

Our attention is on ten traits relevant to resource economy and uptake, competitive ability (or stress tolerance) and reproductive strategy of plants. Seed mass (mg) reflects allocation of energy to a few large vs. many small offspring, and impacts early seedling survival (Moles & Westoby, 2006). Plant height (m) and stem specific density (mg dry mass mm⁻³ fresh volume; hereon SSD) are traits related to light competition, growth rate and long-term viability of the stem (Chave et al., 2009; Moles et al., 2009). Specific leaf area (mm² mg⁻¹; SLA), leaf lifespan (LLS; month), leaf nitrogen and phosphorus content per dry mass (N; mg g⁻¹, P; mg g⁻¹) are traits related to nutrient economy and acquisition and are key components of the leaf economic spectrum (LES; Reich et al., 1999, Wright et al., 2004). SLA represents the potential return rate measured in terms of light capture area per unit of mass investment (Reich et al., 1997; Wright et al., 2004). Leaf lifespan represents the time needed to generate payback on this mass investment (Reich et al., 1992). Leaf N is associated with carboxylation-capacity and is integral to the photosynthetic machinery (Reich et al., 1999; Wright et al., 2004). Leaf P is essential for bioenergetics molecules (e.g. ATP) and is linked to the formation of indispensable nucleic acids and lipid membranes (Reich et al., 1999; Wright et al., 2004). Therefore variation in leaf P and N will be crucial to respiration and photosynthetic capacity, as well as energy generation and storage. Leaf area (mm²) is related to the water and energy balance of a plant and is relevant to light interception. Finally, leaf N and P can also be expressed on an area basis (g m⁻²) reflecting light capture and transaction of energy on an area basis (Wright et al., 2004). Consequently, we also use leaf N per area and leaf P per area, in parallel to their mass-based counterparts.

We obtained spatially explicit trait data for our ten traits and growth form data (woody, non-woody) from the TRY database (www.try-db.org; Kattge et al., 2011; a list of the original data sources is found in Appendix S1). The TRY data subset used in this study includes 16,281 georeferenced records for 15,284 species across all terrestrial biomes sensu Whittaker (1970), from which 9053 and 6231 records were identified as woody and non-woody plants respectively. We standardized the species names to The Plant List (The Plant List, 2013) using Taxonstand (Cayuela & Oksanen, 2014), then we obtained the higher order taxonomy (i.e. family, order, group) for our species with taxonlookup (version 1.0.1; Pennell et al., 2016). The number of individual records in global datasets rapidly decreases as information on more traits is required (Kattge et al., 2011). In our dataset LLS was the trait with the lowest number of records (~0.67% of records have information for this trait), while plant height was the trait with the highest number of records present (~35%; Table S2.1). This limits our ability to assess how traits vary jointly. Thus, we used a hierarchical Bayesian extension of probabilistic matrix factorization to fill in the trait gaps in our dataset (Shan et al., 2012; Fazayeli et al., 2014; Schrodte et al., 2015). This algorithm harnesses the available trait information and the species and higher order taxonomy to fill in the gaps in the trait data. This gap-filling method has been used in other trait analysis at global scale with robust results (see Diaz et al., 2016). We checked the robustness of this method by comparing the trait-trait relationship of only gap-filled data vs. only original data (Table S2.2). We ensure the robustness of our results by re-running our analyses excluding LLS, which is the most gap-filled trait in our dataset (Appendix S3).

Using a map of the Köppen climate zones we assigned the georeferenced plant records to five different climate zones: tropical climate, which includes tropical rainforest, tropical seasonal forest, and savannahs; arid climate, which includes deserts and steppes; temperate climate, which includes temperate forest, temperate rainforest, and Mediterranean vegetation; cold climate, which includes only taiga; polar climate, which includes tundra, alpine and circumpolar zones (Peel et al., 2007).

Estimation of trait networks using precision matrices

Given a set of traits, the precision matrix establishes the undirected statistical dependency among traits. When a coefficient in the precision matrix of traits is different from zero, then a relationship between traits x and y is not due to variation in z (z being a single trait or a set of traits; Horvath, 2011). On the other hand, a value of zero reveals the undirected conditional independency between two traits, given variation in other traits. For instance, if trait x and y are conditionally independent given trait z , that means that traits x and y do not provide information about each other once trait z is considered (Appendix S4). From this it follows that even if traits x and y are highly correlated (Fig. 1c), once trait z is considered any direct relationship between traits x and y would disappear (Fig 1b). Thus, the precision matrix of traits provides the undirected statistical conditional dependency structure among traits for a multivariate (in this case, assumed log-normal) set of traits, which is a graph structure that describes direct probabilistic interactions among traits (Appendix S4; Cai et al., 2012; Ebert-Uphoff & Deng, 2012).

We derived the trait network (i.e. the conditional dependency among traits across organs) for all plants, and then for non-woody and woody plants separately. Next, we derived the trait network for non-woody and woody plants and across the five climate zones defined above. To determine the trait connections among traits across organs we calculated the precision matrix and the confidence interval for each trait-trait interaction in the network. We used a ‘glasso’ algorithm (Graphical Lasso; Friedman et al., 2008) to estimate the precision matrix for each plant group. The glasso algorithm assumes that traits have a multi-variate Gaussian distribution and estimates the precision matrix by minimizing the negative log-likelihood among all plant trait measurements. Thus, prior to analyses we log10 transformed and z-transformed all continuous trait data. The log-likelihood of a Gaussian multivariate distribution is:

$$\begin{aligned}\log f(x) &= \log \left(\frac{1}{\sqrt{(2\pi)^p |\Sigma|}} \exp \left\{ -\frac{1}{2} \mathbf{X}^T \Omega \mathbf{X} \right\} \right) \\ &= -\log \left(\sqrt{(2\pi)^p} \right) - \frac{1}{2} \log |\Sigma| - \frac{1}{2} \text{Tr}(\mathbf{X}^T \mathbf{X} \Omega) \\ &= -\log \left(\sqrt{(2\pi)^p} \right) + \frac{1}{2} \log |\Omega| - \frac{1}{2} \text{Tr}(\hat{\Sigma} \Omega)\end{aligned}$$

where the last equality is from the determinant property: $|\Sigma| = \frac{1}{|\Sigma^{-1}|} = \frac{1}{|\Omega|}$. Thus, ignoring the constants, glasso finds the precision matrix, by solving the following optimization problem:

$$\arg \min_{\Omega} -\log |\Omega| + \text{Tr}(\hat{\Sigma} \Omega) + \lambda \|\Omega\|_1$$

where $\|\Omega\|_1$ is the L_1 norm of Ω .

The glasso algorithm includes a penalty parameter λ , which controls the sparsity level of the precision matrix. If $\lambda = 0$, then all of the elements of the precision matrix are non-zero. In contrast, with a very large value of λ , all the elements become zero. Following Jankova and van de Geer (2015), we accounted for differences in sample size across precision matrices and their effect on λ as follows: $\lambda = 2 \sqrt{(\log p/n)}$, where n refers to the sample size and p denotes the number of nodes, in this case traits (Appendix S4).

Significance of trait-trait connections in the trait network

We tested the significance of the edges for each plant group (i.e. trait-trait interactions) as follows: Given a data sample matrix $X \in \mathfrak{R}^{n \times p}$ and estimated precision matrix $\Omega \in \mathfrak{R}^{p \times p}$, the test statistic is defined as $T = \Omega + \lambda \Omega Z \Omega$ where $Z \in \mathfrak{R}^{p \times p}$ is the sub-gradient of norm $\|\Omega\|_1$, and its variance $\sigma_{ij}^2 = \Omega_{ii} \Omega_{jj} + \Omega_{ij}^2$. Then under this scenario, thresholding T_{ij} at level $\Phi^{-1}(1 - \frac{\alpha}{p(p-1)}) \frac{\hat{\sigma}_{ij}}{\sqrt{n}}$ for all i, j will remove all zero entries (i.e., non-significant trait-trait interactions) with probability $1 - \alpha$ asymptotically (Jankova & van de Geer, 2015).

Specifically, considering the data for a particular group represented in a matrix $X \in \mathfrak{R}^{n \times p}$ which contains the trait information for p traits. First, we fix λ at a particular value and estimate the precision matrix $\Omega \in \mathfrak{R}^{p \times p}$ by running glasso. Next, we estimate the $(1 - \alpha)\%$ asymptotic confidence interval of all the obtained edges as

$$I_{ij} = \left[T_{ij} - \Phi^{-1} \left(1 - \frac{\alpha}{p(p-1)} \right) \frac{\hat{\sigma}_{ij}}{\sqrt{n}}, T_{ij} + \Phi^{-1} \left(1 - \frac{\alpha}{p(p-1)} \right) \frac{\hat{\sigma}_{ij}}{\sqrt{n}} \right]$$

Finally, we removed the edges among traits whose confidence interval contained zero.

Characterization of the trait network

To characterize the differences in the network topology of plant groups obtained through the estimation of precision matrices (described above) we used three metrics: ‘edge density’, ‘modularity’.

‘Edge density’ defines the proportion of present connections among traits out of all possible connections in a network. We used Edge density to describe the connectedness of the trait network. Edge density is used in biological networks, particularly in the study of neural networks, where variation in Edge density has been linked to a compromise between efficiency of connections vs. the cost of connection (Lipson et al., 2002; Alon, 2003). The values for Edge density vary between zero to one. A value of zero represents no connection between the traits of the network. Meanwhile a value of one suggests that all traits in the network are connected to all other traits.

Degree is the number of connections between a trait and other traits, normalized by the total number of unique connections (Horvath, 2011). Thus, we used Degree to quantify the relative importance of the traits for a given trait network. Degree is widely used in biological networks to identify essential characteristics in biological entities such as genes, metabolites, and proteins (Koschftzki & Schreiber, 2008). For example, metabolites with higher Degree may belong to the oldest part of the metabolism, while proteins with higher Degree have been identified as essential, with their removal being lethal to the organism (Koschftzki & Schreiber, 2008). The values for Degree range between zero when a focal trait has no connection to any other trait, to one when a focal trait is connected to all other traits in the network.

‘Modularity’ is the difference between the fraction of connections among traits that fall within a given module (i.e. a module is a subset of traits, that interact more among themselves than with other surrounding traits) minus the same fraction in a null model where connections among traits are randomly distributed (Clauset et al., 2004). Thus, we used Modularity to measure the degree of separation among traits belonging to different modules in terms of fraction of connections that occur within modules (i.e. how connected/disconnected traits across modules are). Modularity helps identify nodes within a network that perform a common function and interact strongly among themselves. Higher Modularity confers an advantage under variable conditions as it provides robustness (Ravasz et al., 2002; Alon, 2003), also providing opportunities for the network to adapt and evolve, as not all components in a modular network are optimally linked (i.e. ability to respond to changing external conditions/internal organization while maintaining normal behaviour; Barabasi & Oltvai, 2004). We used a Spinglass algorithm (see appendix xxx detailed methods) to detect the modules in our networks, as this algorithm accurately detects modules in networks of small size (Number of nodes = 233) and with small or large mixing parameters at the network level (i.e. the summation of external degree of each node over the summation of its total degree; Yang et al., 2016). Modularity values of zero represent networks without a community structure, while non-zero values represent networks that have a community structure.

Significance of trait network metrics

To estimate the statistical significance of the network metrics above we first generated 1000 bootstrapped traits datasets. Using these 1000 databases we calculated the precision matrices for each plant group and climate region (as described above), and calculated the network metrics above for each of the resulting networks. For each resample network we calculated the 95% confidence interval of the difference between groups in a given metric (e.g. edge density, modularity). Following Besag et al. (1995), we corrected for multiple comparison by calculating simultaneous credible interval based on order statistics of the difference in a given network metric (e.g. Degree) among groups.

We ran all analyses under R 3.3.1 (R Core Team, 2016). For calculating the precision matrices we used the camel package (Li et al., 2013), and for calculating the network metrics we used the igraph package (Csardi & Nepusz, 2006).

Results

Among all terrestrial plants, we find a high connectedness between traits across all organs (Edge density = 0.86, 95% CI [0.82-0.89], Fig. 1A). LLS and SSD (both Degree = 1) have significantly higher proportion of connection to other traits than plant height (Degree = 0.86, 95% CI [0.71-0.86]), leaf N (Degree = 0.71, 95% CI [0.71-0.71]), and leaf area (Degree = 0.71, 95% CI [0.71-0.86]). SLA (Degree = 1, 95% CI [0.86-1]) also had a high proportion of connection to other traits, but only significantly higher than plant height and leaf nitrogen (Table S5.1 and 5.2). Seed mass (Degree = 0.86, 95% CI [0.86-0.86]) had higher proportion of connection to other traits, than leaf N. Leaf area (Degree = 0.71, 95% CI [0.71-0.86]) and leaf N mass (Degree = 0.71, 95% CI [0.71-0.71]) were the traits with the fewest connections to other traits (Table S5.1 and 5.2).

Results were generally similar between life forms, but with some significant differences. Both life forms had similar connectedness (Edge density woody = 0.71, 95% CI [0.68-0.75]; Edge density non-woody = 0.61, 95% CI [0.57-0.71]; Fig. 1A). For non-woody species leaf N (Degree = 0.71, 95% CI [0.71-0.86]), leaf area, LLS and leaf P were the traits with the highest number of connections to other traits in the network (Degree = 0.71; 95% CI [0.57-0.86]), while for woody species the traits with the highest proportion of connection to other traits were LLS (Degree = 0.86, 95% CI [0.86-1]), leaf area (Degree = 0.86, 95% CI [0.86-0.86]) and seed mass (Degree = 0.86; 95% CI [0.71-0.86]). In both cases stem related traits were the traits with the fewest connections to other traits in the network (Degree = 0.57). Both life forms had similar levels of modularity, showing higher integration of traits across organs than within organs (Modularity non-woody = 0.10, 95% CI [0.04-0.11]; Modularity woody = 0.06, 95% CI [0.03-0.08]; Fig. 1B). In non-woody species there were two modules, one composed of LLS, leaf N, and SSD, while the other was composed of leaf P, SLA, plant height, leaf area, and seed mass. In woody species we also detected two modules, one formed by leaf area, leaf P and SSD and the other one by seed mass, LLS, SLA, leaf N and plant height.

In both woody and non-woody species, there were connections among all leaf traits, except SLA-LLS in non-woody species. Also, in both groups of species there were connections between plant height-leaf area, plant height-LLS, seed mass-leaf area, seed mass-SLA, seed mass-leaf N, and seed mass-plant height. Connections between LLS-SLA, SSD-leaf area, SSD-leaf P, plant height-leaf N, seed mass-LLS, and seed mass-SSD were only present among woody species. Connections between SSD-leaf N, SSD-LLS, and plant height-leaf P were only present in non-woody species.

Woody and non-woody species by climate region

Woody species show higher proportion of connections among traits (i.e. higher Edge density) in tropical (Edge density = 0.64, 95% CI [0.61-0.68]), temperate (Edge density = 0.68, 95% CI [0.61-0.71]), arid (Edge density = 0.64, 95% CI [0.61-0.78]), and cold environments (Edge density = 0.57, 95% CI [0.53-0.64]), compared to polar ones (Edge density = 0.39, 95% CI [0.29-0.46]; Fig. 2A and Table S5.3 and 5.4). Also, woody species show significantly higher modularity of traits within organs in polar (Modularity = 0.25, 95% [0.16-0.39]) compared to temperate and tropical environments (for both Modularity = 0.08, 95% [0.05-0.11]; Fig. 2B and Table S5.5 and 5.6). Non-woody species show higher proportion of connections between traits in cold (Edge density = 0.61, 95% CI [0.57-0.68]) compared to polar environments (Edge density = 0.43, 95% CI [0.36-0.50]; Fig. 2C and Table S5.3 and 5.4). However, non-woody species also had a higher proportion of connections in temperate (0.57, 95% CI [0.50-0.64]) and cold environments compared to tropical environments (Edge density = 0.36, 95% CI [0.29-0.46]; Fig. 2C and Table S5.3 and 5.4). Non-woody species show non-significant differences in modularity across environments (Fig. 2D and Table S5.5 and 5.6).

Across the different climate regions, woody species always had two modules, while non-woody species had two modules in all climates except polar and tropical where they had three (Table 1). Both growth forms had a module mainly composed of traits related to the LES. In woody species, this consisted of leaf N and leaf P, SLA was part of this module in all climates except tropical, while LLS was part of this module in all climates except cold climate. In non-woody species, it contained SLA, leaf N, and leaf P in all climate regions. The second module was composed of traits related to the reproductive strategy and plant architecture in the case of woody species and mainly plant architecture in the case of non-woody species. In woody species, the core traits in this module were seed mass, plant height and leaf area, and SSD. In non-woody species this module was composed of plant height and leaf area, SSD was part of this module in all climates except polar, and seed mass was part of this module in all climates except tropical and arid. When a third module was present in non-woody species, LLS and either plant height in tropical, or SSD in polar climates composed it (Table 1).

In terms of centrality of traits, as measured by their Degree, for both growth forms LLS was a central trait in temperate (Degree woody = 0.71, 95% CI [0.71-0.86], Degree non-woody = 0.71, 95% CI [0.71-0.86]), and arid climates (Degree woody = 0.86, 95% CI [0.71-0.86], Degree non-woody = 0.57, 95% CI [0.43-0.71]), while seed mass was a central trait in tropical areas (Degree woody = 0.86, 95% CI [0.86-0.86], Degree non-woody = 0.43, 95% CI [0.29-0.71]; Table S5.7-5.10). For non-woody species leaf N was central in all climate zones, except cold, while leaf area was central in all regions except temperate and polar ones (Table S5.9 and 5.10). For woody species LLS was a central trait in all regions except tropical ones (Table S5.7 and 5.8).

Across climate types and both growth forms, connections between leaf N-leaf P, leaf N-SLA, plant height-leaf area and seed mass-leaf area were always present. For non-woody species connections among seed mass-SLA, and seed mass-leaf N were also robust across climate types. Meanwhile, for woody species connections between

SLA-leaf area, leaf P-SLA, LLS-leaf N, SSD-leaf P, seed mass-plant height, and seed mass-SSD were also found across climates regions.

These patterns were largely consistent when we re-ran the analyses using area-based leaf N and P metrics (Appendix S6), or when excluding LLS -the most gap-filled trait in our dataset (Appendix S3). In the area-based results the robustness of connections across climates and growth forms was similar to that observed in mass-based analyses, with the addition of a connection between LLS-SLA, seed mass-leaf N, and seed mass- plant height. In the results excluding LLS we observed the addition of a connection between SSD-leaf area. As in the mass-based results woody species show higher modularity in polar regions, in both area-based metrics (Tables 6.5 and 6.6) and when excluding LLS (Tables 3.5 and 3.6). As in the mass-based results, no unique trait was central (as measured by their Degree) across all climates for both woody and non-woody area-based traits (Tables 6.7-6.10), or when we excluded LLS (Table S3.7-3.10). Across mass-based, area-based results and results excluding LLS we were always able to distinguish two modules, one mainly related to LES traits and another one related to reproductive strategy and plant architecture.

Discussion

In the current study, we used precision matrices of a large global dataset of ten traits that represent all above-ground plant organs (leaf, stem and reproductive) across 16,281 plant species to identify the direct connections (i.e. undirected statistical dependency) that exist across traits within and across organs for this suite of ten traits. We identified emergent characteristics of the trait networks across all land plants at a global scale, as well as across growth forms (woody and non-woody species) and, for the first time (to our knowledge), explicitly accounted for the impact that broad environmental gradients have on the trait network. In doing so, our study builds on - and extends - previous attempts that describe the cross-correlations across several traits and several organs at global scales (Diaz et al., 2016), and efforts that focused on certain vegetation types (Wright et al., 2007) and narrower environmental gradients (Cheng et al., 2015). Important steps forward in this type of analyses will be the incorporation of belowground traits to allow investigation of the connection among LES- root economic spectrum traits across regions (e.g. Liu et al., 2010; Fortunel et al., 2012; Li & Bao, 2015), and that of whole-plant leaf area data to investigate the linkage between lower SLA, longer LLS and whole-plant leaf area (Niinemets, 2010).

Connections among traits across all terrestrial plants

At a global scale we found that, while not all possible trait connections exist or are strong enough for statistical significance, overall land plants have high connectivity across traits and high integration across organs. This supports the idea that matching tissue strategies should be advantageous at the whole-plant scale (Reich, 2014). In terms of traits, LLS, and SSD were the most central traits across land plants. High centrality suggests that a

variable tends to be influential in terms of regulating critical functions or being involved in the regulation of more functions, and therefore of having greater impact on higher level properties, such as fitness (Koschfutzki & Schreiber, 2008). Indeed, LLS is a crucial trait in the resource acquisition strategy of plants, influencing the carbon construction cost and carbon gain of leaves (Reich et al., 1992; Reich et al., 1999), while SSD impacts plant hydraulic and mechanical properties and influences the nutrient, carbon and water economy of stems (Chave et al., 2009). Further, variation in these traits has been shown to impact growth and fecundity (Adler et al., 2014), while also contributing to the structuring of communities (Kunstler et al., 2016) and heavily influencing ecosystem level processes (Reich et al., 1997; Reich, 2014).

Trait connections across growth forms and climate regions

Connections between seed mass-leaf area, leaf N-SLA, leaf P-leaf N, and plant height-leaf area were always present in analyses of growth forms, and of growth forms across different climate types (Table 2). Some of these connections are well known (e.g. leaf N-SLA, leaf P-leaf N), and previous correlation analyses have identified their importance in understanding compromises among traits and their impact on plant function (Reich et al., 1992; Reich et al., 1997; Reich et al., 1999; Wright et al., 2004). Some others are connections that previous correlation analyses suggest have no overall or weak relationships across habitats, and therefore their importance has been downplayed (e.g. seed-mass-leaf area, plant height-leaf area; Table 2). In the first case above - where we detect a connection and previous studies show a strong correlation - evidence suggests that these connections are maintained through selective pressure of biophysical constraints and natural selection (Table 2). For example, an increase in SLA will generally be linked to an increase in leaf N and the concentration of other cytoplasmic molecules (Reich et al., 1997; Meziane & Shipley, 2001). At the same time, natural selection reinforces a strong connection between SLA-leaf N relationship through processes such as herbivory and competition (Reich et al., 1997), limiting the trait space where optimal combinations of these traits occur. In the second case -where conditional dependency between two traits exists, but previous correlation tests suggest a weak relationship- we propose that these trait connections are maintained in the plant phylogeny through neutral or selective processes, but contradictory selective forces across habitats weaken the correlation among these traits. Some studies have reported a triangular relationship between seed mass and leaf area in temperate woody species (i.e. big leaves have big or small seeds, but small leaves only have small seeds; Cornelissen, 1999); other studies report a positive rather than triangular relationships across woody sclerophyll species (Westoby & Wright, 2003), while others across woody tropical species show no relationship among these traits (Wright et al., 2007). In both cases, our analyses suggest that these trait connections are robust after accounting for all other traits, and across climates and growth forms, but the forces maintaining these connections might differ.

Some well-known, strong trait-trait correlations across plants have robust connections globally but not across growth forms or climate regions in our study. For instance, the connection between SLA-LLS is direct in the global data and in woody species but not for non-woody plants (globally). Moreover, this connection was

observed in four of five climate regions for woody plants, but only one of five for non-woody species. This weaker connection in subsets of the global data could suggest that although a strong correlation exists between these two traits even in these subsets, the connection between these traits could be mechanistically mediated through other traits in some cases, or the connection only exists when the absolute range in LLS is large, as is the case for woody plants.

Modularity

Modules are a group of nodes that connect more strongly among themselves than with others, and tend to perform a common function (Alon, 2003). We found that in woody species across climate regions, SLA, leaf N, and leaf P were always together in a module. These three physiological traits are central to the leaf economic spectrum (Wright et al., 2004). Seed mass (a reproductive trait) and plant height and leaf area (architectural traits) formed the core traits in a second module (Table 1). These traits are relevant to plant size, plant-water, and -light relations (Niklas, 1994; Westoby et al., 2002; Moles et al., 2005). Similarly, in non-woody species, leaf N and SLA were always together in one module (leaf P was part of this module, except in tropical climates), while plant height and leaf area were together in a separate module (SSD formed part of this module except in polar climate). Our results suggest that there are at least two core modules across plant growth forms, one module whose function is carbon uptake represented by leaf physiological traits, and a second module whose function is more related to plant size, and plant-light, -water relations represented by seed mass and two architectural traits. The importance of these two modules in the strategies of plants is supported by the global spectrum of plant form and function (Reich et al., 1997; Reich et al., 1999; Diaz et al., 2016).

Modularity evolves in more variable environments (Lipson et al., 2002), as it provides robustness against component failures (Alon, 2003) by lowering the risk of global failures when components in the network experiences external perturbations (Gilarranz et al., 2017). Within woody species modularity is greater in polar than in tropical or temperate climates, consistent with the idea of modularity as providing robustness in stressful or more variable environments. This result also provides some support to the observation that a more modular stem design -more segmented and less hydraulically integrated- occurs under drier conditions (Schenk et al., 2008). Modularity of non-woody species did not significantly differ across climate regions. This pattern corresponds to evidence of non-woody species experiencing weaker environmental filtering (Simova et al., 2018). Taken together, these results suggest that plant growth forms differ in their strategies to cope with environmental conditions, and this is reflected in the connection of traits across the trait network. The change in network modularity in woody species is congruent with the idea that fewer connections between tissues are advantageous in environments that experience more variable conditions, as they may experience greater temporal asynchrony between resources (Bloom et al., 1985).

Conclusion

Previous plant trait studies have focused on understanding the interdependence between pairs of traits or among multiple traits using correlation analyses in combination with ordination techniques (e.g. Table S2.3). These studies have shown that the strength and form of the relationships between traits among co-occurring species varies across environments (Reich et al., 1999; Ackerly, 2004), and among functional groups (Reich et al., 2003), with axes of trait variation shifting, collapsing or arising across environmental gradients and different plant life forms. The analysis presented here suggests that conditional dependencies (i.e. direct from indirect interactions among traits) among multiple traits also shifts across growth forms and climate regions. We found that LLS and SSD are the most central traits globally, but were not always connected within climate zones, especially for non-woody plants; moreover, few trait-trait connections exist robustly across all growth forms and climate gradients (Table 2). Surprisingly, we found that not all strong correlations are direct connections, and some weak correlations are direct connections (Tables S7-S9). Despite the difference in statistical approach with previous research (e.g. Ackerly, 2004; Fortunel et al., 2012; Kramer-Walter et al., 2016), our study supports the existence of two distinguishable dimensions or functional modules across land plants and climate regions. One module is related to physiological leaf traits related to carbon uptake and economy, and another related to reproductive strategy and plant architecture. The findings presented here contribute to the fundamental understanding of dependencies between plant traits and their variation across environmental gradients. The approach taken here represents an important step forward on the collective path to understand the causal links among multiple traits across multiple organs and within and across different climate zones and plant life forms.

Author contributions

HFM, FF, AB, PBR designed the study. FF and AB developed the statistical method. JK, HFM and FF prepared the data. HFM, FF and AD analysed data. HFM and FF wrote the first draft of the manuscript, with all authors contributing to subsequent revisions.