100 YEARS Journal of Ecology

British Ecological Society

Journal of Ecology 2014, **102**, 345–356 doi: 10.1111/1365-2745.12208

SPECIAL FEATURE

THE TREE OF LIFE IN ECOSYSTEMS

Functional distinctiveness of major plant lineages

William K. Cornwell[1,2], Mark Westoby[3], Daniel S. Falster[3], Richard G. FitzJohn[3,4], Brian C. O'Meara[5], Matthew W. Pennell[6], Daniel J. McGlinn[7], Jonathan M. Eastman[6], Angela T. Moles[2], Peter B. Reich[8], David C. Tank[6], Ian J. Wright[3], Lonnie Aarssen[9], Jeremy M. Beaulieu[10], Robert M. Kooyman[3,11], Michelle R. Leishman[3], Eliot T. Miller[12], Ülo Niinemets[13], Jacek Oleksyn[14], Alejandro Ordonez[15], Dana L. Royer[16], Stephen A. Smith[17], Peter F. Stevens[12], Laura Warman[2,18], Peter Wilf[19] and Amy E. Zanne[20,21]*

[1]*Department of Ecological Sciences, Systems Ecology, VU University, de Boelelaan 1085, 1081 HV Amsterdam, The Netherlands;* [2]*Evolution & Ecology Research Centre, School of Biological, Earth and Environmental Sciences, University of New South Wales, Sydney, NSW 2052, Australia;* [3]*Department of Biological Sciences, Macquarie University, Sydney, NSW 2109, Australia;* [4]*Department of Zoology and Biodiversity Research Centre, University of British Columbia, Vancouver, BC V6T1Z4, Canada;* [5]*Department of Ecology and Evolutionary Biology, University of Tennessee, Knoxville, TN 37996, USA;* [6]*Department of Biological Sciences and Institute for Bioinformatics and Evolutionary Studies, University of Idaho, Moscow, ID 83844, USA;* [7]*Department of Biology and the Ecology Center, Utah State University, Logan, UT 84322, USA;* [8]*Department of Forest Resources, University of Minnesota, St. Paul, MN 55108, USA;* [9]*Department of Biology, Queen's University, Kingston, ON K7L 3N6, Canada;* [10]*National Institute for Mathematical & Biological Synthesis, University of Tennessee, Knoxville, TN 37996, USA;* [11]*National Herbarium of New South Wales, Royal Botanic Gardens and Domain Trust, Mrs. Macquaries Rd., Sydney, NSW 2000, Australia;* [12]*Department of Biology, University of Missouri* – St. Louis, St. Louis, MO 63121, USA; [13]*Department of Plant Physiology, Estonian University of Life Sciences, Kreutzwaldi 1, 51014 Tartu, Estonia;* [14]*Polish Academy of Sciences, Institute of Dendrology, 62-035 Kornik, Poland;* [15]*Nelson Institute Center for Climatic Research, Madison, WI 53706, USA;* [16]*Department of Earth and Environmental Sciences, Wesleyan University, Middletown, CT 06459, USA;* [17]*Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, MI 48109, USA;* [18]*Institute of Pacific Islands Forestry, USDA Forest Service, Hilo, HI 96720, USA;* [19]*Department of Geosciences, Pennsylvania State University, University Park, PA 16802, USA;* [20]*Department of Biological Sciences, George Washington University, Washington, DC 20052, USA; and* [21]*Center for Conservation and Sustainable Development, Missouri Botanical Garden, St. Louis, MO 63121, USA*

Summary

1. Plant traits vary widely across species and underpin differences in ecological strategy. Despite centuries of interest, the contributions of different evolutionary lineages to modern-day functional diversity remain poorly quantified.

2. Expanding data bases of plant traits plus rapidly improving phylogenies enable for the first time a data-driven global picture of plant functional diversity across the major clades of higher plants. We mapped five key traits relevant to metabolism, resource competition and reproductive strategy onto a phylogeny across 48324 vascular plant species world-wide, along with climate and biogeographic data. Using a novel metric, we test whether major plant lineages are functionally distinctive. We then highlight the trait–lineage combinations that are most functionally distinctive within the present-day spread of ecological strategies.

3. For some trait–clade combinations, knowing the clade of a species conveys little information to neo- and palaeo-ecologists. In other trait–clade combinations, the clade identity can be highly revealing, especially informative clade–trait combinations include Proteaceae, which is highly distinctive, representing the global slow extreme of the leaf economic spectrum. Magnoliidae and Rosidae contribute large leaf sizes and seed masses and have distinctively warm, wet climatic distributions.

*Correspondence author. E-mail: aezanne@gmail.com

4. *Synthesis.* This analysis provides a shortlist of the most distinctive trait–lineage combinations along with their geographic and climatic context: a global view of extant functional diversity across the tips of the vascular plant phylogeny.

Key-words: determinants of plant community diversity and structure, functional traits, geographic and climatic distributions, Kolmogorov–Smirnov Importance index, leaf nitrogen, leaf size, maximum adult height, phylogenetic tree, seed mass, specific leaf area

Introduction

Traits of plant species quantify how their vegetative and reproductive tissues are structured and how they function. Through the lens of quantitative traits, ecological strategies can be compared across species (Grime 1979). This trait variation is structured spatially and temporally, including across evolutionary and geologic time, continents, climate zones and vegetation types world-wide (Reich, Walters & Ellsworth 1997; Wright *et al.* 2004, 2005; Moles *et al.* 2005; Agrawal 2007). While the role of evolutionary history in shaping contemporary ecology has long been of interest, in recent years, it is increasingly being examined quantitatively (Ackerly & Donoghue 1995; Webb 2000; Cavender-Bares *et al.* 2009; Vamosi *et al.* 2009; Pennell & Harmon 2013).

Studies of functional traits within particular plant groups extend back to the early 1800s (von Humboldt & Bonpland 1807). Since then, researchers have hypothesized about the relative contribution that these different lineages make to the global distribution of plant traits. For example, associated with their root symbiosis, members of the Fabaceae are believed to have increased leaf nitrogen (N) relative to other lineages with implications for the global N cycle (McKey 1994; Vitousek *et al.* 2002; Houlton *et al.* 2008). However, this hypothesis has never been formally tested nor placed in the context of molecular phylogenies or other traits.

Investigating macroecological patterns at large temporal and spatial scales is inherently difficult. The available data are not perfect, being affected by sampling biases with respect to character states, lineages and geographic areas. In addition, we lack a reliable time-calibrated phylogeny for all angiosperms and finely differentiated phylogenies cover a small subset of available trait data.

A more subtle, but perhaps more fundamental, issue is that the assumptions underlying the statistical methods currently available for analysing phylogenetic comparative data become less tenable at larger phylogenetic scales (Felsenstein 2012). For instance, many ecological studies use phylogenies to characterize 'phylogenetic signal', with such approaches as Pagel's λ (Pagel 1999) and Blomberg's K (Blomberg, Garland & Ives 2003) or nonparametric methods such as Mantel (Mantel 1967) and Abouheif (Abouheif & Fairbairn 1997) tests. Phylogenetic signal is often considered to reflect niche conservatism, although these interpretations are questionable because multiple processes can give rise to the same phylogenetic signal (Revell, Harmon & Collar 2008; Pennell & Harmon 2013). Furthermore, both λ and K are based on a simple Brownian motion model, in which rates of evolution are constant across time and across clades. Homogeneous trait evolution is clearly unrealistic for trait data on the scale of all vascular plants (Felsenstein 2012).

More complex models of trait evolution have been developed, which allow for heterogeneity in both rates and models (i.e. not just Brownian motion) across the tree (Butler & King 2004; O'Meara *et al.* 2006). While these may be far more realistic than single rate models, they do not address the question in which we are interested. Most notably, process-based models attempt to characterize the evolution of the trait across the tree. This is distinct – both conceptually and statistically – from our objective, which is to characterize the patterns that result from these evolutionary processes (Uyeda *et al.* 2011).

Ultimately, many processes contribute to produce the observed distribution of traits across extant taxa, including shifts in diversification rates, shifts in the rates of trait evolution or directional evolution within a clade. Because modelling all of these processes simultaneously is not currently tractable, we instead sought to investigate their outcome. More precisely, we aimed to quantify the relative contribution of various clades to the global distribution of important plant functional traits and to highlight the lineages that are exceptionally distinct.

The five traits we examined were specific leaf area (SLA), leaf N, leaf size, maximum adult height and seed mass. These traits capture plant metabolic, competition and reproductive strategies. Specific leaf area (fresh area/dry mass) and leaf nitrogen (N; % mass) content are key components of the 'leaf economics spectrum' (LES; Wright *et al.* 2004). At the 'fast-return' end of the LES are species with high leaf nutrient concentrations, high SLA, short leaf life spans and fast photosynthetic and dark respiration rates. 'Slow return' species have robustly built leaves with long life spans, low nutrient concentrations and slow metabolic rates. Leaf N varies about 24-fold across extant taxa, which although considerable is the least variable trait in our analysis (Table 1). SLA varies almost 500-fold across species. Leaf size varies more than a million-fold, is correlated with twig size along a scaling spectrum known as Corner's Rules (Westoby *et al.* 2002) and has important consequences for the leaf energy budget via leaf temperature. Canopy height at maturation varies more than 10 000-fold and captures an aspect of light competition and life span of the main stem (Falster & Westoby 2003; Moles *et al.* 2009). Seed mass varies more than 10^{11}-fold and reflects allocation to few large versus many small offspring for a given amount of energy and has implications for dispersal and early seedling survival (Moles *et al.* 2005).

Molecular systematics has brought new clarity to evolutionary relationships among plants, sometimes reinforcing historical hypotheses based on morphology and other times bringing together unexpected relatives. In this study, we made use of

Table 1. Global summary statistics for the five traits in this analysis

	Units	Number of species	Median	Minimum	Maximum
Leaf N per mass	%	4481	1.8	0.3	6.0
Specific leaf area	$cm^2\ g^{-1}$	6868	137	3	1441
Maximum height	m	21 626	2.0	0.001	112
Leaf size	cm^2	8751	16.3	1.9×10^{-3}	2.6×10^3
Seed mass		31 937	2.3 mg	0.03 μg	21 kg

these new phylogenetic trees to tackle several goals: using a new phylogenetic method, we identify the most distinctive lineages for each trait. In some cases, distinctive lineage–trait combinations confirm particular anecdotal hypotheses, and in other cases, the lineage–trait combinations are completely novel. We then examine the context of these clades, including the age of the lineage and the spatial and climatic distribution of extant species.

A new method for identifying functionally distinctive lineages

To identify influential lineages, we devised a new metric based on classic statistics for comparing frequency distributions (Kolmogorov 1933), the Kolmogorov–Smirnov Importance index (KSI). The method is designed to identify lineages that significantly alter the distribution of trait values observed in current day taxa. Without these lineages, the modern distribution of functional traits would be very different. The key features of the new method are that it examines trait information only from extant taxa and balances the dual influences of species richness and functional differentiation when identifying distinctive clades.

Given a distribution of traits among extant taxa, and a phylogenetic tree for these taxa, the KSI measures the functional distinctiveness of all possible clades on the tree. As this method does not require branch lengths, all that is required is a topology for the tree. For each node on the tree, we compare the frequency distribution of trait values for species within the clade descended from that node to the distribution for all other species in the tree (Fig. 1a–c), using a nonparametric two-sample Kolmogorov–Smirnov test. Such a test asks how likely is it that the groups in each comparison came from the same distribution, and as such, does this clade contribute a different range of trait values to that observed elsewhere in the tree? We are not interested in testing whether a particular clade shows a significant difference in trait distribution, but rather in using the measure of difference calculated in the Kolmogorov–Smirnov significance test as a way to rank the functional distinctiveness of different clades.

The difference between the trait distributions within and outside any given clade can be compared based on the maximum difference in their cumulative distributions $F_i(x)$ and $F_j(x)$ (Fig. 1):

$$D_{i,j} = \max|F_i(x) - F_j(x)| \qquad \text{eqn 1}$$

The likelihood that the two trait distributions come from the same underlying distribution is then a monotonic function of our KSI:

$$I_{i,j} = \sqrt{((n_i * n_j)/(n_i + n_j)) * D_{i,j}}, \qquad \text{eqn 2}$$

where n_i and n_j are the number of species in the two groups (Kolmogorov 1933). With increasing values of I, it becomes increasingly unlikely that species in the clade of interest come from the same underlying distribution as species in the neighbourhood. We then compare I for all clades. The clade with the largest value of I has the highest probability of differing in its trait distribution to the rest of the tree (Fig. 1) and as such is the most functionally distinctive.

One advantage of the KSI is that it naturally balances species richness and functional differentiation. By combining the raw difference in cumulative distribution functions with the sample size of the clades in question, this analysis is designed to identify clades that are both very unusual with respect to their trait values and contain many species. Moreover, the Kolmogorov–Smirnov test is able to distinguish not only differences in mean trait value, but also changes in variance and skewness of trait distributions among comparison groups.

To determine the top 5 most distinctive clades, we used a hierarchical algorithm, such that after the first clade was selected, the comparison group for each clade was redefined. Clades within the first-selected clade were compared with other species subtended from the first-selected node (Fig. 1). Clades outside the first-selected clade were compared with other species outside the first-selected clade (Fig. 1). The process was repeated, with new values of $D_{i,i}$ and $I_{i,j}$ calculated based on the refined neighbourhoods of each clade; this is analogous to the algorithm used by Alfaro *et al.* (2009). The KSI code is open source and is available as an R package (http://github.com/richfitz/ksi).

Materials and methods

TRAIT DATA

We assembled a data base for 48 324 species for five traits: SLA (leaf area/leaf mass), leaf nitrogen (N) concentration, seed mass, maximum height and leaf size (see sample sizes for each trait in Table 1). These data are a compilation of separate research initiatives focusing on specific traits; data were gathered directly from researchers leading those individual efforts and/or the literature (see Supporting Information for further details about individual trait data sets).

SPECIES AND LINEAGE NAMES

To bring species binomials to a common taxonomy across data sets, names were matched against the accepted names in the Plant List

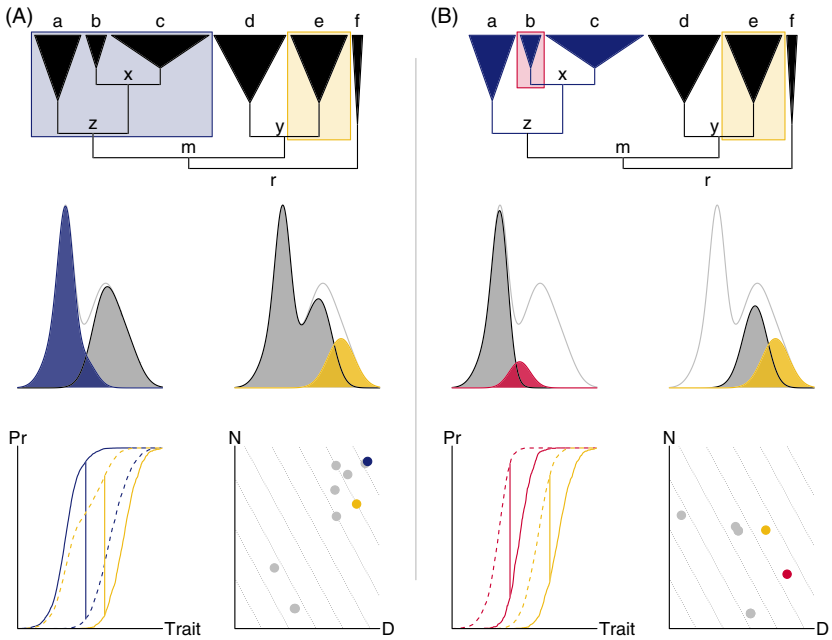

Fig. 1. How the Kolmogorov–Smirnov Importance (KSI) index is used to identify exceptionally distinctive lineages for a given trait. On the first pass (Panel A), the phylogeny is broken at all possible nodes; for each node, the trait distribution for all descendants is compared against the distribution of all remaining taxa in the tree using a Kolmogorov–Smirnov (KS) test. For example, for node 'z', the trait distribution for members of clades a–c (blue region on tree and blue distribution in middle plot) is compared against the distribution of clades d–f (grey distribution in middle plot), with the grey line representing the combined distribution of the two groups. The Kolmogorov–Smirnov method computes the greatest difference (*D*) between the empirical cumulative distribution functions for these two groups' trait distributions (bottom left plot; height of blue vertical bar is test statistic for node 'z', with the dashed line representing the grey distribution from the top left panel). This is repeated for all nodes and clades in the tree – for this six-taxon tree, there would be 10 possible comparisons, and a comparison for clade 'e' versus clades {a–d,f} is indicated in yellow. Following Kolmogorov–Smirnov, the statistic is weighted by the sample size of the two groups (*N*, see text for details) and the clade with the highest product of *N* and *D* is taken as the 'most distinctive' (bottom right plot; diagonal lines indicate isoclines of *N***D*). In this case, clade 'z' is the most distinctive. The KSI index is applied recursively to the tree; in the second and subsequent rounds, rather than comparing a clade's distribution to the whole tree, we compare it to its neighbourhood (Panel B). The previous most distinct clade was 'z', dividing the tree into two regions (blue vs. black colouring on the phylogeny). Clades are then only compared against regions of the tree with the same colour – so 'b' (in red) is compared against {a,c} and 'e' (in yellow) is compared against {d,f} (c.f. panel A). For each comparison, we compute the Kolmogorov–Smirnov statistic *D* (bottom left plot) and the sample size weighting to identify the next most distinctive clade. In this second round, we find that clade 'e' is the most distinctive, followed by clade 'b'.

(http://www.theplantlist.org/). Any binomials not found in this list were then matched against the International Plant Names Index (IPNI; http://www.ipni.org/) and Tropicos (http://www.tropicos.org/); potential synonymy in binomials arising from the three lists was investigated using the Plant List tools. Binomials remaining unmatched were compared first with the Plant List and then IPNI with an approximate matching ('grepping') algorithm. For binomials with accepted generic names but unmatched binomials, we first searched for specific epithet misspellings within the genus and then broadened the search to all plants to check whether the generic name was incorrect. We then searched for unmatched generic names. We found that including the species epithet in the approximate matching algorithm with the full list of binomials improved determination of the correct genus.

With the steps above and a strict approximate grepping threshold (roughly corresponding to one letter substitution or a gender error in the specific epithet) and when there was only one match returned, the false positive rate was low (< 1%) and could be automated. When the threshold was relaxed to look for names that still did not match, the false positive rate rose to unacceptable levels. For these species and for those that returned multiple matches, we examined and made potential substitutions on an individual case basis.

With regard to higher lineage names above the level of genus, we followed APG III (Chase *et al.* 2009) for ordinal level and below and followed Cantino *et al.* (2007) and Soltis *et al.* (2011) for above ordinal level as applicable.

PHYLOGENETIC TREE

Because KSI does not require branch length information, we did not seek to build a phylogeny with accurate branch lengths. Instead, we sought to construct a phylogeny that had consensus support and included all species in our trait data set. Species were mapped onto a consensus phylogeny (Chase *et al.* 2009) using Phylomatic (Webb & Donoghue 2005) with further sub-family-level resolution from the Angiosperm Phylogeny Website (Stevens 2001). For the main analysis, we used a tree that included all species within our trait data base. For the sampling bias analysis, we built a tree that included all the accepted names in the Plant List (http://www.theplantlist.org/).

CLIMATE AND GEOGRAPHY

Species binomials were queried against the Global Biodiversity Information Facility (GBIF; www.gbif.org/) data base to extract georeference points, which were used to determine species' climate niches by mapping points against an interpolated climate grid. Overlap with the

five trait data bases ranged from 42% to 48% coverage. We used a series of criteria to filter the GBIF records (see Supporting Information for details). All georeference locations were queried against Worldclim (Hijmans *et al.* 2004) 5-arc-min resolution data products to determine point location estimates of mean annual temperature (MAT), annual precipitation (AP) and seasonal standard deviation of MAT and AP. For each binomial, we calculated the median for climate variables and latitude. These median climate estimates were analysed using the same KSI methods as the trait values (see exact algorithm in Supporting Information).

ANALYSES

We applied our new method for determining functional distinctiveness based on KSI to the data base and trees described above. The KSI test is insensitive to data transformations, but for presentation purposes (e.g. Fig. 2), species mean trait values were logged.

Because of the nested nature of phylogenies and strong tree imbalance in many parts of the phylogeny, there is variation in certainty of the precise membership in the lineage that contributes most to extant diversity using KSI. We identified a population of nested lineages of interest. We then selected the lineage with the highest test statistic as our focal distinctive lineage in the main text; we also show the population of lineages (Fig. S1 in Supporting Information).

Results

We identified the five top-ranked lineages for each trait (Fig. 2; Table 2; *P*-values are not relevant at this scale, but for all reported lineages are $< 10^{-10}$). Another way to consider this ranking is as follows. If one were to divide the probability density function of traits into an increasing number of evolutionary groups, this ranking sequentially finds the lineages for which there is the most statistical evidence. For example, for maximum height, the available evidence points to the first split at monocots-minus-Acorales, the second at euasterids, and so forth. The KSI analysis balances the extremeness of the trait values with the number of taxa in the lineage. Because of this balance, the clades selected here include some small lineages with extreme trait values and large clades, which may have less extreme trait values.

Each lineage selected by the KSI (Tables 2 and 3; Figs 2 and 3) differed in its own way, although some common patterns emerged. In some cases, extreme trait values were almost entirely found within one clade. For example, globally low leaf N species were almost entirely within the Proteaceae (Figs 2 and S2). In contrast, high leaf N was found within many clades that also contain less extreme trait values. Similarly, monocots-minus-Acorales and euasterids make up the majority of the world's short species (Fig. 3). Tall species come from a number of clades that also include shorter species.

In principle, the trait distribution within clades identified by the KSI analysis may vary from that of the neighbourhood in a number of ways, including mean, variance, skewness or kurtosis. However, in practice, for the nodes at the top of the list identified here, the primary difference in the distribution was a shift in the means (Fig. 3). In the case of maximum height, the global bimodal distribution was sometimes in contrast to a unimodal distribution within a specific clade (Fig. 3). There were also cases of strong shifts in spread, without a shift in the mean, but these examples were not within the top 5 nodes identified for these five traits. Shifts in variance deserve subsequent attention with the new class of methods for examining heterogeneity in evolutionary rates.

These top lineages are typically part of a population of nested high-ranking nodes for that trait. With five traits examined for their top-5 most distinct lineages, we had the potential to recover 25 distinct lineages. However, because several lineages (or nested lineages) were top-5-ranked for more than one trait, only 12 notable clades were identified (Tables 2 and 3; Fig. 2). In some cases, these noteworthy trait–lineage combinations provide insight into geographic and climatic distributions of species; in others, the lineage, while maintaining distinctive trait values, is found globally (Table 2; Fig. S3).

Discussion

Below, we examine different trait spectra in turn. We group together SLA and leaf N, which are components of the LES (Wright *et al.* 2004), and seed mass and maximum height, which are moderately correlated ($r = 0.64$; in part due to links between maternal investment in offspring and duration of competitive growth; Falster, Moles & Westoby 2008). In some cases, the quantification here makes long-standing but anecdotal knowledge of lineages more precise. In other cases, the results are entirely surprising, especially where KSI analysis identified previously unnamed clades.

LEAF ECONOMIC SPECTRUM

The leaf economic spectrum influences carbon and nutrient cycling rates across the world. Three distantly related lineages, Proteaceae (1600 species (Stevens 2001)), Ericaceae-plus-closely related families (4445 species (Stevens 2001)) and Acrogymnospermae (947 species (Stevens 2001)), all contributed to the slow end of the LES (Figs 2 and 3; Proteaceae and Ericaceae, respectively, ranked 1 and 2 for both leaf N and SLA and Acrogymnospermae ranked 4 for leaf N and 5 for SLA). These three lineages are common in low nutrient soils across the world although Proteaceae is absent from the temperate part of the Northern Hemisphere (rank 1 for Southern Hemisphere distribution; Fig. S3). In the absence of these lineages, especially Proteaceae, the span of the global LES would be markedly narrower. All three clades have noteworthy adaptations to extract resources in low-fertility contexts: some Proteaceae have specialized cluster roots to remove P from occluded forms in old soils (Lambers *et al.* 2008) while root symbioses with Ascomycetes and Basidiomycetes are common within both Ericaceae and Acrogymnospermae (Cornelissen *et al.* 2001). Sarraceniaceae, a family closely related to Ericaceae and included in the selected clade, has a carnivorous nutrient acquisition strategy.

Rosids-minus-Vitales (70 000 species (Stevens 2001)) were also identified as having low SLA (rank 3) although they lack

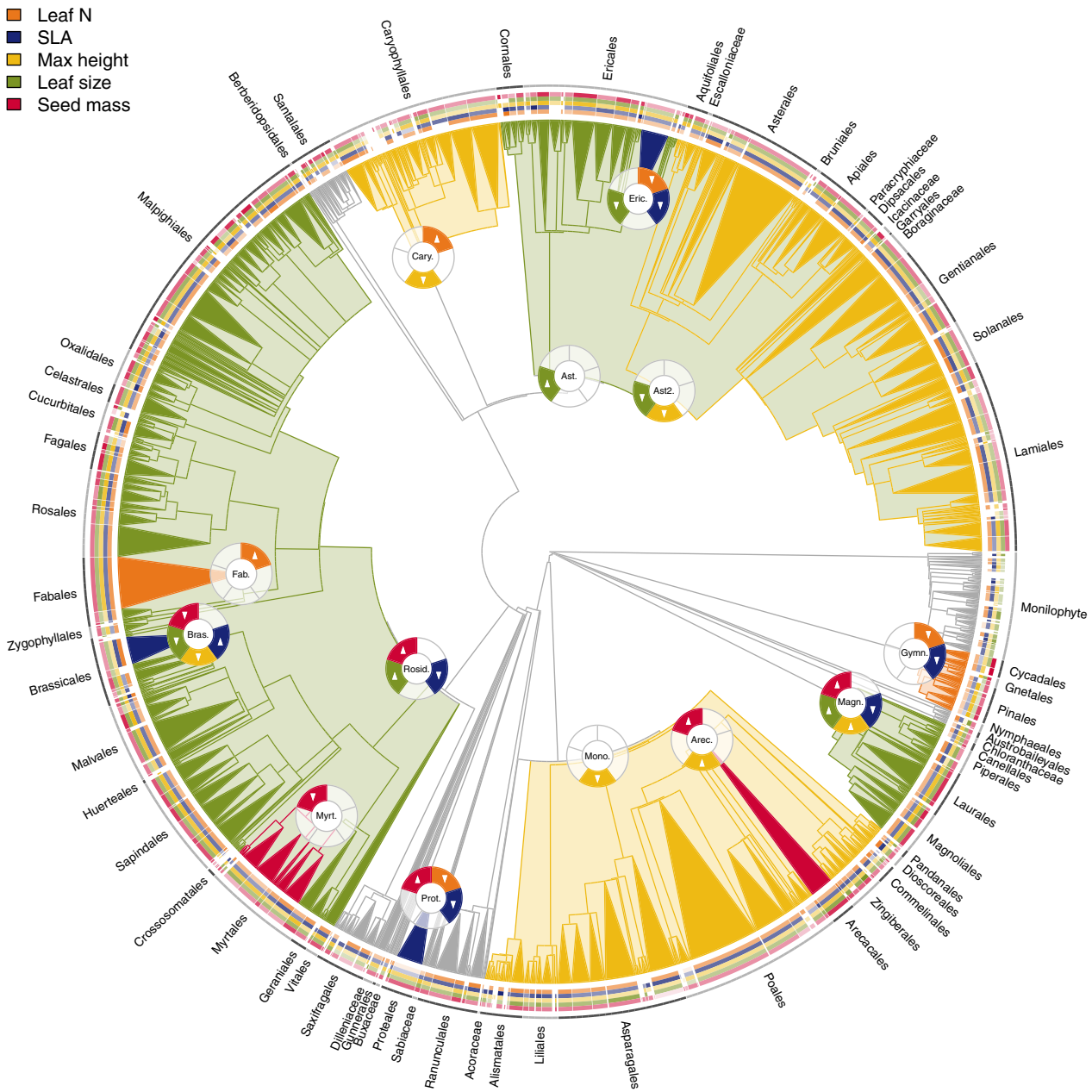

Fig. 2. The most functionally distinctive lineages for each of five traits considered across vascular plants world-wide. Distinctive clades are indicated by the small pie charts. Segment colours and clade colours indicate the particular trait for which that lineage was selected with arrows indicating the direction of change. The background colour visually indicates the nested structure of the clades. When the uncertainty in selected lineages for multiple traits overlap, we only highlight one lineage (see Fig. S1 and Table 2). The raw trait values for all five traits are shown in concentric circles around the tree with the intensity of the colour corresponding to family level means for traits. For example, Proteaceae (lower left of Fig. 1) has very low intensity orange and blue segments indicating its position at the extreme slow end of the leaf economic spectrum (low SLA and leaf N). Orders are shown around the outside of the tree. Branch lengths in this tree are for graphical purposes only and do not reflect estimated dates. SLA, specific leaf area.

a single identified root specialization that can be associated with a slow return position on the LES. Despite having descendants spread across the globe, the vast majority of the rosid lineage is in equatorial (rank 2), warm (rank 1 for MAT), wet (rank 2 for annual precipitation) places. Within this typical rosid, climatic context dwells an exception: Brassicaceae (3710 species (Stevens 2001)), which is a young clade noteworthy for its high SLA (rank 4) often growing in temperate, disturbed environments (Franzke *et al.* 2010) with low annual precipitation (rank 4) and low precipitation seasonality (rank 8).

We expected both the extreme high and extreme low values for each trait would be comprised of distinctive lineages. However, this was not the case. For the leaf economic spectrum, there were many more distinctive clades at the slow end compared with the fast end (Figs 2, 3 and S2). Ranking species by their position on the LES (using their position on a SMA axis for leaf N and SLA), 45 of the lowest 50 species

Table 2. The distinctive lineage–trait combinations, including number of species per lineage, that contribute to extant diversity in five functional traits (maximum height, seed mass, leaf N, leaf size and specific leaf area = SLA), climate (MT, mean annual temperature; AP, annual precipitation; TS, temperature seasonality; and PS, precipitation seasonality) and latitude. As substantial overlap occurs in the lineages selected across traits, 12 total are described. Within each trait, numbers indicate the relative rank in the global search algorithm. When clades for a given trait also rank within the top 10 for a different trait, climate or latitude, these ranks are also reported. Arrows indicate the direction of the shift. Row lineage names give an approximate location within the phylogeny; when actual lineages recovered are different from the name in the first column, specific lineage membership is reported within the footnotes

Lineage	Number of species*	Maximum height	Seed mass	Leaf N	Leaf size	SLA	Climate	Latitude
Asteridae	92 861	2† (↓)			3 (↑), 5† (↓)			
Rosidae	70 000		1 (↑)		2 (↑)	3‡ (↓)	MT 1 (↑), AP 2 (↑)	2 equatorial
Monocotyledonae-minus-Acorales	60 096	1 (↓)						
Fabaceae	19 500			3 (↑)			PS 7 (↓)	
Caryophyllales	11 510	4 (↓)		5 (↑)			AP 3 (↓)	
Myrtales-minus-Combretaceae	10 527		5 (↓)					6§ southern distribution
Magnoliidae	9900	7¶ (↑)	2¶ (↑)		1 (↑)	6¶ (↓)	MT 2 (↑), AP 1 (↑), TS 3 (↓)	
Ericaceae (plus nearby families)	4445			2** (↓)	4 (↓)	2†† (↓)		
Brassicaceae	3710	5‡‡ (↓)	6 (↓)			4 (↑)	AP 4 (↓), PS 8 (↓)	
Arecaceae	2361	3 (↑)	3 (↑)				MT 6 (↑), AP 6 (↑)	
Proteaceae	1600		4 (↑)	1 (↓)		1 (↓)		1 southern distribution
Acrogymnospermae	947			4§§ (↓)		5 (↓)		

*Stevens (2001), Wang *et al.* (2009)
†Gentianidae (=euasterids).
‡Rosidae-minus-Vitales.
§Myrtaceae.
¶Magnoliales + Laurales.
**Ericaceae + Cyrillaceae + Clethraceae + Sarraceniaceae + Roridulaceae + Actinidiaceae + Diapensiaceae + Styracaceae + Symblocaceae + Theaceae.
††Ericaceae + Cyrillaceae + Clethraceae + Sarraceniaceae + Roridulaceae + Actinidiaceae.
‡‡Brassicaceae + Cleomaceae.
§§Pinales-minus-Pinaceae.

were in the Proteaceae (including the 29 most extreme species). Conversely, the highest 50 species were from 28 distantly related families with Fabaceae the most common (appearing only six times). Furthermore, unlike Proteaceae, those 28 families also contained many species with moderate, as well as high trait values. In other words, the global slow end of the LES was composed of only a few distinct clades, while the fast end was composed of species from many different evolutionary lineages, each of which also contain species with less extreme trait values.

Two further notable clades with respect to the LES were the Fabaceae and the Caryophyllales. Our results confirmed the hypothesis that Fabaceae have distinctively high N leaves (McKey 1994), and our analysis also revealed an important nuance: Fabaceae sit towards the fast end of the LES but were shifted orthogonally from the main LES spectrum, having distinctly high leaf N (rank 3) at a given SLA (Fig. S2). This may be aligned with a wide array of successional strategies in the family (Menge, DeNoyer & Lichstein 2010). High leaf N is presumably associated with the N-fixing rhizobial symbioses in a majority of Fabaceae species; however, other related clades within rosids also have N-fixing symbioses (Soltis *et al.* 1995) but did not stand out as having increased leaf N in this analysis. For Caryophyllales, high leaf N (rank 5) is part of a suite of specialized water relations traits, also including succulence and C_4 or CAM photosynthesis. Caryophyllales are especially abundant in arid (rank 3 for annual precipitation; Fig. S3), saline and disturbed environments or have a carnivorous or epiphytic habit.

SEED MASS AND MAXIMUM HEIGHT

Across all species, seed mass and maximum height were moderately linked with large seeds associated with large stature adults, as found previously (Moles & Westoby 2006; Falster, Moles & Westoby 2008). In some notable clades, large seeds were associated with tall stature, while in other clades, only one of the two traits was remarkable in its distribution. Significant associations between these traits may be driven by the particular ecological contexts in which the lineages are distributed.

Monocots-minus-Acorales (59 300 species) ranked 1 for their short canopy height. Monocots are well recognized as dominating biomass in semi-arid environments (Woodward & Lomas 2004), but they are distributed across the globe in almost all of the world's climates. Interestingly, they were not exceptional for any trait or climate variable other than their short stature. Arecaceae (palms; 2361 species (Stevens 2001)) within monocots-minus-Acorales are a prominent exception to the rule of short monocot herbs, though their 'pseudo-woody'

Table 3. Relative rank and components of KSI (see equations 1 and 2) for the five most distinctive lineages that contribute to extant diversity in five functional traits (maximum height, seed mass, leaf N, leaf size and specific leaf area = SLA). Plus signs (+) indicate the identified lineage includes the set of named lineages (see Stevens 2001 and onwards for named lineage descriptions); minus signs (−) indicate the identified clade is a subset of the preceding lineage, excluding the following clade. In other words, Myrtales-Combretaceae lineage includes all of the species in Myrtales except those within Combretaceae. These calculations all follow Kolmogorov (1933). KSI is derived from D (the distinctiveness) and $\sqrt{((n_i * n_j)/(n_i + n_j))}$, the weighting for number of species in the comparison. Because the comparison groups change as the tree is partitioned (see Fig. 1), KSI is not monotonic with the rank

Lineage	Rank	KSI (I from eqn 2)	D (from eqn 1)	$n1$ (comparison group s.r.)	$n2$ (target group s.r.)	$\sqrt{((n_i * n_j)/(n_i + n_j))}$ (see eqn 2)
Maximum height						
Nartheciidae	1	22.82	0.39	13 446	4583	58.46
Gentianidae	2	18.37	0.35	9563	3883	52.55
Arecaceae	3	17.28	0.80	4062	521	21.49
Caryophyllales	4	12.41	0.49	8882	681	25.15
Brassicaceae + Cleomaceae	5	11.49	0.71	8610	272	16.24
Seed mass						
Rosidae	1	23.28	0.31	18 791	8104	75.25
Magnoliales + Laurales	2	14.70	0.84	18 483	308	17.41
Arecaceae	3	14.45	0.91	18 230	253	15.80
Proteaceae	4	14.48	0.66	17 735	495	21.94
Myrtales-Combretaceae	5	13.64	0.53	7376	728	25.74
Leaf N						
Proteaceae	1	10.39	0.86	3512	151	12.03
Ericaceae + Cyrillaceae + Clethraceae + Sarraceniaceae + Roridulaceae + Actinidiaceae + Diapensiaceae + Styracaceae + Symplocaceae + Theaceae	2	6.55	0.55	3366	146	11.83
Fabaceae	3	6.03	0.39	3108	258	15.43
Pinales-Pinaceae	4	4.51	0.71	3067	41	6.36
Caryophyllales	5	4.14	0.33	2900	167	12.57
Leaf size						
Magnoliidae	1	8.57	0.42	6917	449	20.53
Rosidae	2	9.22	0.22	3984	2933	41.10
Asteridae	3	7.20	0.23	1692	2292	31.20
Ericaceae	4	5.08	0.41	2128	164	12.34
Gentianidae	5	5.65	0.32	384	1744	17.74
SLA						
Proteaceae	1	9.89	0.78	5655	166	12.70
Ericaceae + Cyrillaceae + Clethraceae + Sarraceniaceae + Roridulaceae + Actinidiaceae	2	5.94	0.45	5475	180	13.20
Rosidae-Vitales	3	4.80	0.14	3505	1970	35.51
Brassicaceae	4	5.17	0.50	1857	113	10.32
Acrogymnospermae	5	4.47	0.63	3454	51	7.09

KSI, Kolmogorov–Smirnov Importance.

growth is not achieved via secondary thickening. Palms ranked 3 for both tall canopy height and large seeds (Linkies *et al.* 2010). Arecaceae largely inhabit wet, warm (rank 6 for both annual precipitation and temperature) tropical and subtropical areas, where tall stature and large seeds are common features of canopy dominants from many clades. Besides Arecaceae, two other lineages have both large seeds and primarily a tropical distribution: the Magnolales-plus-Laurales (rank 2 for seed mass; Fig. S3) and the Rosidae (rank 1 for seed mass). Many shade-tolerant tropical species possess large seeds, thought to be important for seedling establishment when carbon fixation is limiting (Moles & Westoby 2006).

Within the generally large-seeded rosids, Myrtales-minus-Combretaceae (10 527 species (Stevens 2001)) were highlighted for small seeds (rank 5). This previously unremarked upon clade, predominantly from the Southern Hemisphere (rank 6), includes shrubs and trees from open vegetation (Myrtaceae), tropical weeds and epiphytes (Melastomataceae), and herbs (most of Onagraceae–Lythraceae). While small seeds are spread throughout this clade, and regeneration after disturbance is a common theme, biome and growth forms within the clade are remarkably diverse.

Proteaceae also had large seed mass (rank 4), although the functional role of these large seeds may be driven by retention of P in poor soils experiencing frequent fires rather than low light. This Proteaceae strategy occurs within the close geographic neighbourhood of the Myrtales-minus-Combretaceae strategy of small seeds, suggesting that within a given environment different lineages may have consistently different successful strategies for regeneration. Asterids (rank 2) and

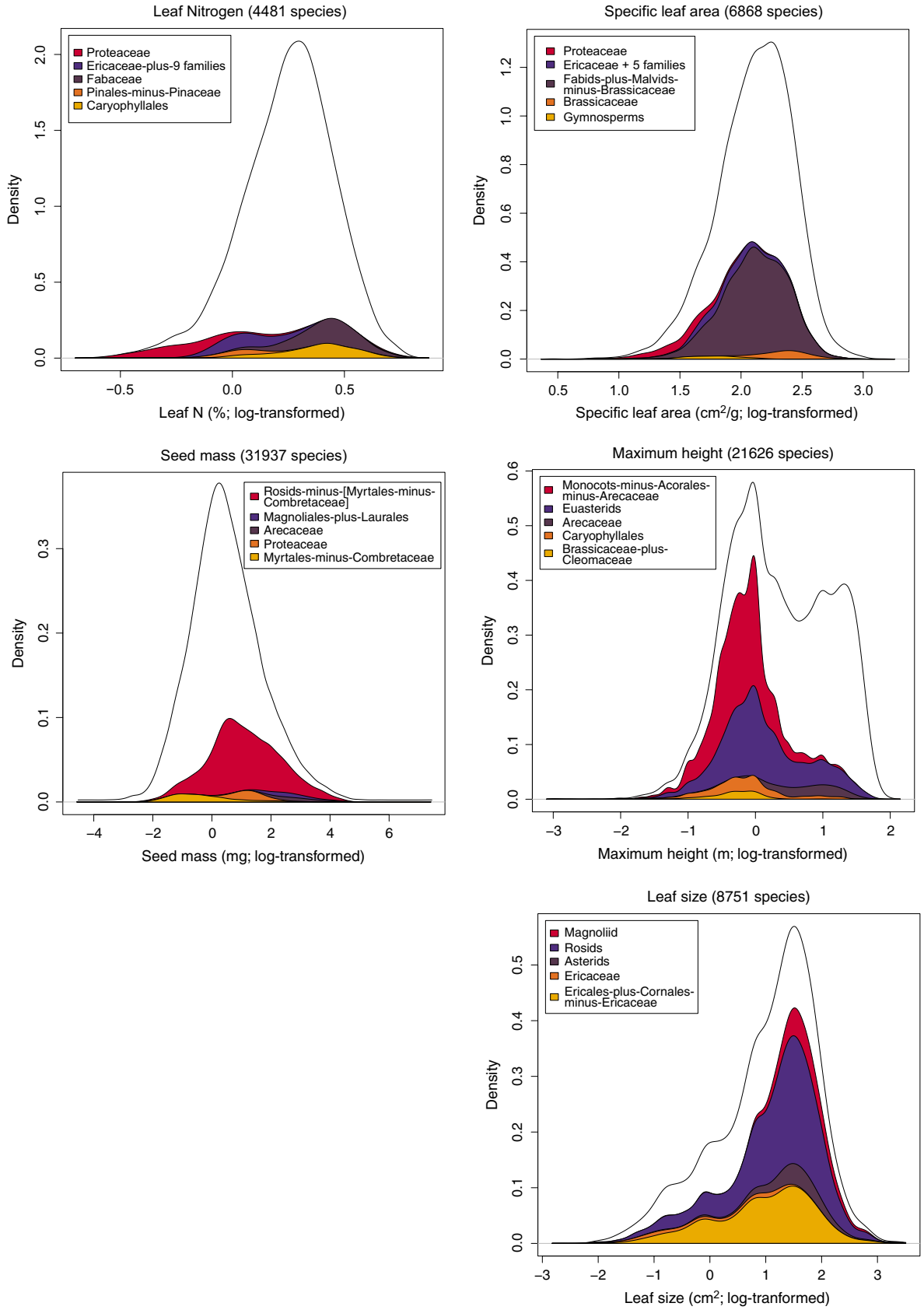

Fig. 3. Density plots for the top 5 selected lineages for each of the five traits with the total area of each colour scaled proportionally to clade richness within the trait data set. The black line denotes the density plot for the traits of all plants in the data set. The density plots are stacked vertically; the areas for each clade should not be read as overlapping except that all the identified clades form a portion of global diversity. Note: the extreme distribution of Proteaceae with respect to both leaf N and SLA (and see Fig. S2 for the bivariate plot). SLA, specific leaf area.

the Caryophyllales (rank 4) showed a tendency towards short height. Part of this pattern may be explained by a distribution in seasonal environments: the lamids (within the asterids) are associated with high seasonality in rainfall. Brassicaceae also had short stature (rank 5), as well as small seeds (rank 6).

As discussed above for the leaf economic spectrum, there is strong directional asymmetry with regard to how clades were functionally distinctive for maximum height. Most of the world's short species come from two clades: Gentianidae (=euasterids) and Monocots-minus-Acorales. These two clades comprised 66% of the shortest quartile of species (shorter than 0.6 m), but only 23% of the world's tallest quartile of species (taller than 10 m). In contrast, the globally tall species were from many different evolutionary lineages, each of which also contained species with less extreme trait values. For example, the tallest species in the world (*Sequoia sempervirens*, *Psuedotsuga menziesii* and *Eucalyptus regnans*) have numerous short cousins.

LEAF SIZE

As an overall pattern, the large-leaved clades were primarily found in tropical environments: Magnoliidae (9900 species (Stevens 2001); rank 1 for leaf size) are disproportionately low-latitude (Fig. S3), high-precipitation (rank 1) and high-temperature (rank 2) specialists, with limited seasonal variation in temperature (rank 3). While speciose and successful within tropical forests, Magnoliidae have largely failed to flourish in other biomes. Rosids also had large leaf sizes (rank 2) relative to other clades and a warm wet distribution as discussed above.

The distribution of leaf size in the asterid clade was complex, with KSI identifying three nested lineages (Fig. 2). The first was the asterid clade itself with large leaves (rank 3). Within that clade were two lineages with small leaves: the campanulids-plus-lamids, also known as the Gentianidae (rank 5) and Ericaceae (rank 4). With those two clades removed, the result was that the Ericales-plus-Cornales-minus-Ericaceae (coloured green in Fig. 2) comprised the remaining large-leaved species.

SAMPLING BIAS IN FUNCTIONAL TRAIT DATA BASES

This is a synthesis of trait data from 48 324 species, 17% of documented diversity (http://www.theplantlist.org/). That said, we note that this is a much smaller number of species than that for which genetic data are available in GenBank (84 838 species), or for which geographical observations are available in GBIF. Further progress remains to be made in quantifying and synthesizing trait data, in particular. Additionally, all of these large data sets are non-random samples of the entire vascular plant phylogeny (see discussion within Smith *et al.* (2011) for biased sampling within GenBank). For functional traits, we suggest future research efforts prioritize measuring traits in under-sampled clades. To this end, we used our approach to identify and rank clades in the vascular plant phylogeny where functional measurements are notably lagging behind taxonomic knowledge. We again used the KSI method to find the top 5 disproportionately under-sampled clades; they are Orchidaceae (1% sampled), Gentianales (7%; especially Rubiaceae within Gentianales: 6%), Gesneriaceae (1%), Bromeliaceae (2%) and Araceae (2%). These speciose clades are often in the canopies or understories of tropical forests across the globe, representing an under-sampled part of global functional diversity, at least from a quantitative perspective. While our current analysis is by far the largest to date, and the clade–trait combinations we identify are the most important given current data, future research on under-sampled clades may well shift the ranking in Table 3 and discover new functionally distinct lineages.

Conclusions

The striking array of form and function among plant species has been discussed through two hundred years of ecological, evolutionary and systematic scrutiny (von Humboldt & Bonpland 1807). Ecologists increasingly are incorporating evolutionary history as part of their conceptual framework for understanding present-day ecology (Ackerly & Donoghue 1995; Webb 2000; Cavender-Bares *et al.* 2009; Vamosi *et al.* 2009; Pennell & Harmon 2013). There is an emerging picture that most species tend to resemble their close relatives, but of course this is not always the case: rates of trait evolution can be rapid in some parts of a phylogeny and slow in others (O'Meara *et al.* 2006; Ackerly 2009). Furthermore, species within certain clades may be similar to each other but as a group highly convergent with other distantly related groups. We show here that for some trait–clade combinations, knowledge of a species' lineage conveys little information to neo- and palaeo-ecologists. In other trait–clade combinations, the lineage identity can be highly revealing. For example, knowing a species is a member of the rosid clade does not reveal anything about its maximum height, but a rosid is more likely than the average plant to have large seeds and large leaves and to be from close to the equator.

Our results are a mix of quantitatively testing anecdotal knowledge and discovering new patterns, via novel, rigorous quantitative methods. Proteaceae have long been regarded as an emblematic Southern Hemisphere clade (Fig. S3), but here, we show that they uniquely extend the slow return end of the global leaf economic spectrum. Caryophyllales have adopted diverse adaptations to aridity (Fig. S3), but are unified by high leaf N concentrations and short stature. Species within the Myrtales-minus-Combretaceae are diverse in habitat but have small seeds in common. Differences in functional trait space coincide with clades inhabiting predominantly tropical (magnoliids; Fig. S3, rosids, Arecaceae) or temperate (Brassicaceae) regions. These globally distinctive clade–trait combinations provide key puzzle pieces to the jigsaw of modern plant functional diversity.

Acknowledgements

We thank members of the Tempo and Mode of Plant Trait Evolution working group for contributing to project development and manuscript comments, including Michael Donoghue, Erika Edwards, Ginger Jui, Hafiz Maherali, Risa Sargent and Elisabeth Wheeler. We acknowledge LEDA, the NZ flora, the Australian Biodiversity Resources Study (ABRS) /Australian Flora online, Ian Dickie, Kenwin Liu, Bonnie Jacob, Liz Law, Ginger Jui, Nathan Swenson and the NCEAS Neotropical rain forest communities working group for contributions to the underlying data sets. We also thank Andrea Hahn and Tim Robertson from GBIF who kindly provided species' georeference points and Alan Paton and Nicola Nicolson who kindly provided IPNI lists. This work was supported by NESCent, Macquarie University Genes to Geoscience Research Centre through the working group, and NWO. D.J.M. was supported by a grant from the NSF DEB-0953694.

Author contributions

WKC, AEZ and SAS led the working groups. WKC, MW and AEZ wrote the original manuscript and coordinated analyses and further writing. DSF, RGF, BCO and MWP were major quantitative contributors especially with new methods development, analyses, graphics and writing. DJM coordinated the spatial and climatic data. JME, DJM, ATM, PBR, DCT and IJW were large contributors through initial ideas development, methods, data set curation, analyses and writing. LA, JMB, RMK, MRL, ETM, UN, JO, AO, DLR, SAS, PFS, LW and PW contributed data sets and discussions, and read drafts.

References

Abouheif, E. & Fairbairn, D.J. (1997) A comparative analysis of allometry for sexual size dimorphism: assessing Rensch's rule. *American Naturalist*, **149**, 540–562.

Ackerly, D.D. (2009) Conservatism and diversification of plant functional traits: evolutionary rates versus phylogenetic signal. *Proceedings of the National Academy of Sciences of the USA*, **106**(Suppl 2), 19699–19706.

Ackerly, D.D. & Donoghue, M.J. (1995) Phylogeny and ecology reconsidered. *Journal of Ecology*, **83**, 730–733.

Agrawal, A.A. (2007) Macroevolution of plant defense strategies. *Trends in Ecology & Evolution*, **22**, 103–109.

Alfaro, M.E., Santini, F., Brock, C., Alamillo, H., Dornburg, A., Rabosky, D.L., Carnevale, G. & Harmon, L.J. (2009) Nine exceptional radiations plus high turnover explain species diversity in jawed vertebrates. *Proceedings of the National Academy of Sciences of the USA*, **106**, 13410–13414.

Blomberg, S.P., Garland, J.R. & Ives, A.R. (2003) Testing for phylogenetic signal in comparative data: behavioral traits are more labile. *Evolution*, **57**, 717–745.

Butler, M.A. & King, A.A. (2004) Phylogenetic comparative analysis: a modeling approach for adaptive evolution. *The American Naturalist*, **164**, 683–695.

Cantino, P.D., Doyle, J.A., Graham, S.W., Judd, W.S., Olmstead, R.G., Soltis, D.E., Soltis, P.S. & Donoghue, M.J. (2007) Towards a phylogenetic nomenclature of tracheophyta. *Taxon*, **56**, 1E–44E.

Cavender-Bares, J., Kozak, K.H., Fine, P.V. & Kembel, S.W. (2009) The merging of community ecology and phylogenetic biology. *Ecology Letters*, **12**, 693–715.

Chase, M.W., Fay, M.F., Reveal, J.L., Soltis, D.E., Soltis, P.S., Anderberg, A.A., Moore, M.J., Olmstead, R.G. & Rudall, P.J. (2009) An update of the angiosperm phylogeny group classification for the orders and families of flowering plants: APG III. *Botanical Journal of the Linnean Society*, **161**, 105–121.

Cornelissen, J.H.C., Aerts, R., Cerabolini, B., Werger, M.J.A. & van der Heijden, M.G.A. (2001) Carbon cycling traits of plant species are linked with mycorrhizal strategy. *Oecologia*, **129**, 611–619.

Falster, D.S., Moles, A.T. & Westoby, M. (2008) A general model for the scaling of offspring size and adult size. *The American Naturalist*, **172**, 299–317.

Falster, D.S. & Westoby, M. (2003) Plant height and evolutionary games. *Trends in Ecology & Evolution*, **18**, 337–343.

Felsenstein, J. (2012) A comparative method for both discrete and continuous characters using the threshold model. *The American Naturalist*, **179**, 145–156.

Franzke, A., Lysak, M.A., Al-Shehbaz, I.A., Koch, M.A. & Mummenhoff, K. (2010) Cabbage family affairs: the evolutionary history of brassicaceae. *Trends in Plant Science*, **16**, 1360–1365.

Grime, P.J. (1979) *Plant Strategies and Vegetation Processes*. Wiley, New York.

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. & Jarvis, A. (2004) The worldclim interpolated global terrestrial climate surfaces. Version 13. Available at http://biogeo.berkeley.edu/.

Houlton, B.Z., Wang, Y.P., Vitousek, P.M. & Field, C.B. (2008) A unifying framework for dinitrogen fixation in the terrestrial biosphere. *Nature*, **454**, 327–330.

von Humboldt, A. & Bonpland, A. (1807) *Essai sur la géographie des plantes*. Levarault Schoell, Paris.

Kolmogorov, A.N. (1933) *Grundbegriffe der wahrscheinlichkeitsrechnung*. Julius Springer, Berlin.

Lambers, H., Raven, J.A., Shaver, G.R. & Smith, S.E. (2008) Plant nutrient-acquisition strategies change with soil age. *Trends in Ecology & Evolution*, **23**, 95–103.

Linkies, A., Graeber, K., Knight, C. & Leubner-Metzger, G. (2010) The evolution of seeds. *New Phytologist*, **186**, 817–831.

Mantel, N. (1967) The detection of disease clustering and a generalized regression approach. *Cancer Research*, **27**, 209–220.

McKey, D. (1994) Legumes and nitrogen: the evolutionary ecology of a nitrogen-demanding lifestyle. *Advances in Legume Systematics*, **5**, 211–228.

Menge, D.N., DeNoyer, J.L. & Lichstein, J.W. (2010) Phylogenetic constraints do not explain the rarity of nitrogen-fixing trees in late-successional temperate forests. *PLoS ONE*, **5**, e12056.

Moles, A.T. & Westoby, M. (2006) Seed size and plant strategy across the whole life cycle. *Oikos*, **113**, 91–105.

Moles, A.T., Ackerly, D.D., Webb, C.O., Tweddle, J.C., Dickie, J.B. & Westoby, M. (2005) A brief history of seed size. *Science*, **307**, 576–580.

Moles, A.T., Warton, D.I., Warman, L., Swenson, N.G., Laffan, S.W., Zanne, A.E., Pitman, A., Hemmings, F.A. & Leishman, M.R. (2009) Global patterns in plant height. *Journal of Ecology*, **97**, 923–932.

O'Meara, B.C., Ané, C., Sanderson, M.J. & Wainwright, P.C. (2006) Testing for different rates of continuous trait evolution using likelihood. *Evolution*, **60**, 922–933.

Pagel, M. (1999) Inferring the historical patterns of biological evolution. *Nature*, **401**, 877–884.

Pennell, M.W. & Harmon, L.J. (2013) An integrative view of phylogenetic comparative methods: connections to population genetics, community ecology, and paleobiology. *Annals of the New York Academy of Sciences*, **1289**, 90–105.

Reich, P.B., Walters, M.B. & Ellsworth, D.S. (1997) From tropics to tundra: global convergence in plant functioning. *Proceedings of the National Academy of Sciences of the USA*, **94**, 13730.

Revell, L.J., Harmon, L.J. & Collar, D.C. (2008) Phylogenetic signal, evolutionary process, and rate. *Systematic Biology*, **57**, 591–601.

Smith, S.A., Beaulieu, J.M., Stamatakis, A. & Donoghue, M.J. (2011) Understanding angiosperm diversification using small and large phylogenetic trees. *American Journal of Botany*, **98**, 404–414.

Soltis, D.E., Soltis, P.S., Morgan, D.R., Swensen, S.M., Mullin, B.C., Dowd, J.M. & Martin, P.G. (1995) Chloroplast gene sequence data suggest a single origin of the predisposition for symbiotic nitrogen fixation in angiosperms. *Proceedings of the National Academy of Sciences of the USA*, **92**, 2647.

Soltis, D.E., Smith, S.A., Cellinese, N., Wurdack, K.J., Tank, D.C., Brockington, S.F., Refulio-Rodriguez, N.F., Walker, J.B., Moore, M.J. & Carlsward, B.S. (2011) Angiosperm phylogeny: 17 genes, 640 taxa. *American Journal of Botany*, **98**, 704–730.

Stevens, P.F. (2001) Angiosperm phylogeny website. Version 9, June 2008 [and more or less continuously updated since]. Available at http://www.Mobot. Org/MOBOT/research/APweb.

The Plant List (2010) Available at http://www.theplantlist.org/.

Uyeda, J.C., Hansen, T.F., Arnold, S.J. & Pienaar, J. (2011) The million-year wait for macroevolutionary bursts. *Proceedings of the National Academy of Sciences USA*, **108**, 15908–15913.

Vamosi, S.M., Heard, S.B., Vamosi, J.C. & Webb, C.O. (2009) Emerging patterns in the comparative analysis of phylogenetic community structure. *Molecular Ecology*, **18**, 572–592.

Vitousek, P.M., Cassman, K., Cleveland, C., Crews, T., Field, C.B., Grimm, N.B., Howarth, R.W., Marino, R., Martinelli, L. & Rastetter, E.B. (2002) Towards an ecological understanding of biological nitrogen fixation. *Biogeochemistry*, **57**, 1–45.

Wang, H., Moore, M.J., Soltis, P.S., Bell, C.D., Brockington, S.F., Alexandre, R., Davis, C.C., Latvis, M., Manchester, S.R., & Soltis, D.E. (2009) Rosid radiation and the rapid rise of angiosperm-dominated forests. *Proceedings of the National Academy of Sciences*, **106**, 3853–3858.

Webb, C.O. (2000) Exploring the phylogenetic structure of ecological communities: an example for rain forest trees. *American Naturalist*, **156**, 145–155.

Webb, C.O. & Donoghue, M.J. (2005) Phylomatic: tree assembly for applied phylogenetics. *Molecular Ecology Notes*, **5**, 181–183.

Westoby, M., Falster, D.S., Moles, A.T., Vesk, P.A. & Wright, I.J. (2002) Plant ecological strategies: some leading dimensions of variation between species. *Annual Review Ecology and Systematics*, **33**, 125–159.

Woodward, F.I. & Lomas, M.R. (2004) Vegetation dynamics – simulating responses to climatic change. *Biological Reviews*, **79**, 643–670.

Wright, I.J., Reich, P.B., Westoby, M., Ackerly, D.D., Baruch, Z., Bongers, F. *et al.* (2004) The worldwide leaf economics spectrum. *Nature*, **428**, 821–827.

Wright, I.J., Reich, P.B., Cornelissen, J.H.C., Falster, D.S., Groom, P.K., Hikosaka, K., Lee, W., Lusk, C.H., Niinemets, Ü. & Oleksyn, J. (2005) Modulation of leaf economic traits and trait relationships by climate. *Global Ecology and Biogeography*, **14**, 411–422.

Received 30 June 2013; accepted 9 December 2013
Handling Editor: Amy Austin

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Data S1. Supporting methods, figures and references for data used in analyses.