Traits, trees and taxa: global dimensions of biodiversity in mammals

Shan Huang*, Patrick R. Stephens and John L. Gittleman

Odum School of Ecology, University of Georgia, Athens, GA 30605, USA

Measures of biodiversity encompass variation along several dimensions such as species richness (SR), phylogenetic diversity (PD) and functional/trait diversity (TD). At the global scale, it is widely recognized that SR and PD are strongly correlated, but the extent to which either tends to capture variation in TD is unclear. Here, we assess relationships among PD, SR and TD for a number of traits both across clades and regional assemblages of mammals. We also contrast results using two different measures of TD, trait variance and a new measure we refer to as trait bin filling (the number of orders of magnitude of variation that contain at least one species). When TD is defined as trait variance, PD is a much stronger correlate of TD than SR across clades, consistent with hypotheses about the conservation value of PD. However, when TD is defined as bin filling, PD and SR show similar correlations with TD across clades and space. We also investigate potential losses of SR, PD and TD if species that are currently threatened were to go extinct, and find that threatened PD is often a similar predictor of threatened TD as SR.

Keywords: biodiversity; species richness; phylogenetic diversity; trait diversity; mammals

1. INTRODUCTION

Quantifying biodiversity is a central issue in ecology, biogeography and conservation biology [1–3]. Considered broadly, biodiversity encompasses variation in the composition and characteristics of species assemblages (e.g. clades and communities). The diversity of assemblages can vary along several dimensions such as species richness (SR, taxonomic diversity sensu [4]), genetic diversity and trait diversity (TD) or functional diversity [1,4–6]. Of these, TD has been the least studied even though it is potentially one of the most important aspects of biodiversity from a conservation perspective. For example, communities with high TD and trait redundancy among species might be more resilient to perturbation [6,7]. Conversely, when species with distinctive ecological characteristics are lost from a local community, their extinction can have a disproportionate effect on the species that remain, and alter or reduce ecosystem function [8–11]. Unfortunately, despite the availability of many sophisticated mathematical algorithms that have been developed for calculating TD of species (reviewed in Pavoine & Bonsall [3]), in most real-world cases, it is not feasible to quantify TD for regional assemblages owing to our limited knowledge of species’ distributions and traits. Even when a complete species list is available in many cases, the traits of all but a few well-studied species are poorly known. For example, in mammals, perhaps the best-studied class of vertebrates, estimates of adult body mass are available in the literature for only approximately 60 per cent of species [12]. Other characteristics such as litter size and population density have been quantified for even fewer species [12].

Recently, there has been much interest in quantifying the phylogenetic diversity (PD) of species assemblages [13,14], defined as the summed branch lengths of a phylogeny of the species in an assemblage [15]. PD reflects both the number and the evolutionary distinctiveness of species in an assemblage, and thus can potentially act as a ‘silver bullet’ ensoncing several dimensions of biodiversity [13,16–18]. We know from previous studies that PD is strongly correlated with SR in many systems [2,4,19–21], due in large part to the way PD is calculated (i.e. when you add more species to a tree, the summed branch lengths of the tree are bound to increase). When PD is calculated from a tree with branch lengths that reflect sequence divergence, it can be presumed to also capture genetic diversity above the species level [15,22]. Whether PD generally captures variation in TD remains an open question. Several studies have investigated the relationship between PD and trait variance in experimental plant systems [23–25], but with mixed results. A global study of mammals considered the relationships among SR, PD and a composite measure of functional diversity [5], and showed all three variables to be strongly correlated; however, this analysis included only four traits, and the study did not assess variation in the strength of correlations for individual traits.

If PD is, in general, correlated with TD for a wide range of traits, then it could serve as an important proxy for a dimension of biodiversity that is difficult to quantify otherwise. Knowledge of the phylogenetic structure of a community could also provide useful information such as which species might have unique ecological roles based on their phylogenetic distinctiveness. However, SR is easier to quantify than PD, the latter of which requires not only knowledge of which species occur in an assemblage but also their evolutionary relationships. If PD is not a stronger correlate of TD than SR, then the extra effort required to quantify PD (i.e. constructing a phylogeny) would not be warranted when the primary goal is to estimate expected TD. To date, this issue has rarely been investigated empirically. Fritz & Purvis [26] found...
that if threatened mammal species were to go extinct, the geographical patterns of PD losses across the world’s ecoregions [27] would not be consistent with the losses of TD measured as adult body mass variance. However, the spatial scale they used (i.e. entire ecoregions) is arguably coarse for many conservation considerations, and they did not assess the relationship between total TD and total PD, examine traits other than mass, or consider measures of TD other than variance [26].

Here, we investigate the relationship between TD, PD and SR for a variety of traits in terrestrial mammals for two types of assemblage: across phylogenetic clades and across geographical regions. We first investigate the relationship between PD and TD across clades, because if no correlation occurs at the clade level it seems unlikely that such a relationship would hold between the PD and TD of communities (which can be thought of as somewhat random subsamples of full mammalian clades). We quantify TD in both individual-level traits, such as litter size, as well as in population and species-level traits such as population density. We perform analyses to address three questions: (i) Is SR and PD correlated with TD across clades? (ii) Is PD generally a stronger correlate of TD than SR? Finally, (iii) does sample size or variation in phylogenetic signal explain differences in the strength of correlations between TD and PD or SR? Intuitively, the strength of phylogenetic signal in a trait could affect the degree to which PD is a better predictor than SR, because traits with lower signal exhibit variation that is more independent of phylogeny. Additionally, we expect that sample size for an individual trait might affect the observed patterns of TD. Because the amount of data available for mammals varies dramatically for different traits [12], low proportions of mammal species in a clade or a community that a trait has been sampled for could add ‘noise’ to estimates of correlations between TD and PD (or SR). We further conduct global spatial analyses using two traits (body mass and geographical range size) that have particularly large sample sizes to see (iv) whether PD (or SR) can be a good representative of TD across geographical regions; (v) whether the areas of highest PD and TD show geographical correspondence (i.e. are ‘hotspots’ [28,29] of PD and TD in mammals congruent) and (vi) from a conservation perspective, whether the number of threatened species and amount of PD loss in a region predict the amount of TD that stands to be lost.

2. METHODS

(a) Defining trait diversity

We here use the term ‘trait diversity’ in preference to ‘functional diversity’ in acknowledgement of the fact that the biological traits we considered may or may not directly reflect the ecosystem function of a species (e.g. mice and shrews have very different ecological roles, though they are both small, have high reproductive rates, and can reach high population densities). Among existing measures of TD (reviewed in [3,30]), trait variance [26,31] and composite measures of functional diversity [4,32,33] are perhaps the most widely used in recent broad-scale studies. A previous global study of mammals [5] used a composite measure of trait variation based on dendrograms constructed using the unweighted pair group method with arithmetic averages (i.e. UPGMA). However, the summed branch lengths of a UPGMA tree will inevitably be correlated with the summed branch lengths of a phylogeny for the same number of species because both are sensitive to SR, even when UPGMA trees are generated using completely random trait data (see the electronic supplementary material, figure S1). Thus, their method is not well suited for testing for a general relationship between TD and PD. Other ways to examine composite TD include principal components and a multidimensional measure of functional richness such as FRv or FRIm (sensu [30]). However, the number of species that we could include using these methods would diminish rapidly as we included more traits because they all require complete case analysis (see the electronic supplementary material, figure S2). Both to maximize sample sizes and to avoid using a measure of TD that is auto-correlated with PD and SR, we focused on the variance of individual traits in this study.

While trait variance is a fairly standard measure of TD, it does not necessarily capture TD in the sense of niche filling. A measure of niche filling that has been used in previous studies is functional range, the range of species trait values in a local community compared with the range of species trait values in a region or globally [34]. While this does capture variation in the breadth of niche space occupied communities, it does not fully reflect ‘niche filling’ in that a community with two species (e.g. a community consisting of a mouse species and an elephant species) could potentially be considered as diverse as a community containing many species with a greater variety of trait values. We therefore propose a new measure of TD that we refer to as trait bin filling, which more closely reflects TD in the classical community ecological sense of the filling of semi-discrete niches (e.g. MacArthur’s ‘broken stick’ model of niche partitioning; [35]) than either trait variance or other measures considering functional range [34]. Bin filling refers to the number of orders of magnitude of a given trait occupied by at least one species in a clade or community. For example, body mass spans roughly eight orders of magnitude in mammals, and so an assemblage of mammal species can potentially fill up to eight mass bins. If a community consisted of one large mammal of body mass $1.4 \times 10^6$ g and 10 small mammals of between 1.0 $\times 10^2$ and 9.9 $\times 10^2$ g, it would have two mass bins filled. In some cases where it was desirable to have larger numbers of bins, such as analyses of traits that span few orders of magnitude of variation in mammals and when producing global maps of bin filling, we used natural log bins instead of log10 bins.

We acknowledge that mammals similar in certain traits can fulfil very different ecological roles in a community (e.g. rodents versus insectivores). However, it at least stands to reason that species that differ by an order of magnitude in a trait such as mass or litter size have important ecological differences. Furthermore, the bin filling definition of TD is potentially easier to measure accurately than variance or many other TD measures. Most species traits exhibit some degree of allometric scaling [36]. In general, medium-to-large-bodied species that occur in communities are known, whereas most unknown species will tend to be small-bodied. Owing to the greater overall number of small-bodied species in most communities and the potential for such species to reach extremely high abundance, at least a few small-bodied species will generally have been recorded for any given community [37]. By definition, as long as one
species is reported falling into a given trait bin, the bin is considered filled. Thus, estimates of bin filling are potentially more robust to incomplete species sampling than other TD estimates.

(b) **Clade-level analyses**

Twelve traits extracted from PanTHERIA [12] and IUCN (see below) were used for clade level analyses. Included traits span a wide variety of sample sizes from 535 to 4642 species (table 1; electronic supplementary material, figure S2) and a wide range of phylogenetic dependence (Blomberg’s $K$ [38,39]), from $K=0.079$ to 1.855 (table 1). Despite the wide variation in signal observed, all traits showed statistically significant signal ($\text{sensu } [38]$) when compared to randomized data ($p<0.01$ in every case). Analyses were conducted for both trait variance and trait bins using a super-tree of all extant mammals [10]. We analysed correlations using Spearman’s rank-order test, and compared Spearman’s $\rho$-values to see which of SR and PD was more tightly correlated with TD (see the electronic supplementary material for more details). All the clade analyses were conducted in R v. 2.12.2 [40], phylogenetic analyses were performed using the package APE v. 2.7 [41].

(c) **Geographical analyses**

Geographical analyses focused on two traits for which we had particularly good species-level sampling, adult body mass and geographic range area (table 1; electronic supplementary material, figure S3). Both traits are considered important in ecology, evolution, and conservation biology [8,26,37,42]. Geographical distribution data and threat status data of all terrestrial mammals were obtained from the IUCN database (www.iucnredlist.org) and were matched to the taxonomy [43] used in the phylogenetic and trait data. SR, PD and the two measures of TD using two traits, and the hypothetical reduction of them given extinction of threatened species were calculated for assemblages occurring in each $100 \times 100 \text{km}^2$ grid cell of an equal-area projection of the world’s land surface. We analysed spatial correlations using Spearman’s rank (see the electronic supplementary material for more details). All the geographical data and results were processed in ArcGIS and all the spatial analyses were conducted in R [40].

### 3. RESULTS

(a) **Clade-level analyses**

Variance of eight of 12 traits showed significant correlations with PD, whereas variances of only three showed significant correlations with SR (see the electronic supplementary material, table S1). The observed correlations (Spearman’s $\rho$) were generally higher between PD and TD compared with those between SR and TD (seven of eight traits that showed significant correlations). However, both PD and SR were similar predictors of bin filling. Five of 12 traits showed significant correlations for both PD and SR, whereas one additional trait showed a correlation with SR alone and one with PD alone. Whether bin filling shows a stronger correlation with PD or SR varied among traits. Blomberg’s $K$ did not show a significant correlation with the strength of correlations observed across traits ($n=12$; trait variance versus PD: $\rho=0.164$, $p=0.609$; trait variance versus SR: $\rho=0.217$, $p=0.499$; trait bins versus PD: $\rho=-0.245$, $p=0.444$; trait bins versus SR: $\rho=-0.168$, $p=0.604$) or the degree to which correlations with PD and SR differed ($n=12$; trait variance: $\rho=0.147$, $p=0.651$; trait bins: $p=0.111$, $p=0.733$).

There was no correlation between sample size and the strength of correlations observed for trait variance ($n=12$; PD: $r^2=0.021$, $p=0.591$; SR: $r^2=0.193$, $p=0.088$) or trait bins ($n=12$; PD: $r^2=0.046$, $p=0.427$; SR: $r^2=0.017$, $p=0.589$). Simulations showed that trait variance shows no autocorrelation with PD or SR (see the electronic supplementary material, figure S4). However, bin filling did show significant autocorrelation with both PD and SR, with either variable explaining roughly 45 per cent of the variation in the number of ‘bins’ filled with draws from a random lognormal distribution (see the electronic supplementary material, figure S5).

(b) **Geographical analyses**

All three dimensions of biodiversity showed complex patterns of variation across the globe (figure 1; electronic supplementary material, figures S6 and S7). When trait variance was used as our measure of TD, extremely weak but (generally) statistically significant global correlations between the TD and the PD or SR of regional assemblages were observed (see the electronic supplementary material, table S2). By contrast, strong correlations were observed at a continental scale (see the electronic supplementary material, table S3). PD loss appeared to be a good predictor of TD loss in terms of trait variance, especially for body mass, at both global and continental scales; the number of threatened species (i.e. threatened SR) was also correlated with TD loss though to a lower degree for body mass (see the electronic supplementary material, table S2 and S3 and for additional discussion of these results). We further found that TD was significantly correlated with PD and SR in cells where residual PD was either very high (top five percentile), such as those in the eastern to southern Africa and India, or very low (bottom five percentile), such as those in western north America and the Amazon area (see the electronic supplementary material, table S4 and figure S6). However, in these areas, PD did not appear to be a much stronger predictor of TD than SR.

---

**Table 1. Traits included in clade analyses.**

<table>
<thead>
<tr>
<th>Trait</th>
<th>Sample size</th>
<th>Blomberg’s $K$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Geographical range area (km$^2$)</td>
<td>4762</td>
<td>0.079</td>
</tr>
<tr>
<td>Adult body mass (g)</td>
<td>3468</td>
<td>2.146</td>
</tr>
<tr>
<td>Litter size (no. of offspring)</td>
<td>2478</td>
<td>0.504</td>
</tr>
<tr>
<td>Adult head-body length (mm)</td>
<td>1910</td>
<td>1.554</td>
</tr>
<tr>
<td>Gestation length (day)</td>
<td>1359</td>
<td>3.172</td>
</tr>
<tr>
<td>Population density</td>
<td>950</td>
<td>0.286</td>
</tr>
<tr>
<td>(number km$^{-2}$)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adult forearm length (mm)</td>
<td>892</td>
<td>1.671</td>
</tr>
<tr>
<td>Litters per year (no. of litters)</td>
<td>889</td>
<td>0.358</td>
</tr>
<tr>
<td>Home range area (km$^2$)</td>
<td>703</td>
<td>0.419</td>
</tr>
<tr>
<td>Teat number (no. of teats)</td>
<td>628</td>
<td>0.454</td>
</tr>
<tr>
<td>Basal metabolic rate (ml O$_2$ h$^{-1}$)</td>
<td>571</td>
<td>0.834</td>
</tr>
<tr>
<td>Mass-specific metabolic rate (ml O$_2$ h$^{-1}$/g)</td>
<td>571</td>
<td>0.888</td>
</tr>
</tbody>
</table>
When the number of trait bins filled by species in a grid cell was used as our measure of TD, PD and SR showed strong positive correlations with TD (see the electronic supplementary material, table S2 and S3). On the basis of comparison of the correlations, PD had a stronger predictive power than SR for the TD of geographical range area but not adult body mass. SR and PD losses also showed significant correlations with the number of bins that would be ‘emptied out’ if all threatened species were to go extinct, though the correlations observed were much stronger for mass than range area (see the electronic supplementary material, table S2 and S3 and for additional discussion of these results). At a global scale, there was relatively little overlap between the areas of highest diversity for PD and the two measures of TD. Areas of high PD and trait variance were generally outside CI hotspots. However, trait bin ‘hotspots’ were generally well captured by the CI hotspots (see the electronic supplementary material, figure S8).

4. DISCUSSION

Whether or not PD is a ‘silver bullet’ that captures TD in addition to SR depends on the definition of TD used and the scale of analysis. At the clade level, when TD is defined as trait variance, models based on PD are generally a better fit than models based on SR (see the electronic supplementary material, table S1). However, when TD was defined as bin filling, which of the two was a better fit varied from one trait to the next and both showed similar overall performance (i.e. when SR showed a significant correlation PD generally also did and vice versa). In general, PD is at worst a similar fit as SR to models of TD for clades, and it is often a better fit. When both PD and SR are available for mammal clades PD is generally to be preferred as a surrogate for expected TD, though the difference between the two measures is sometimes modest (i.e. $\Delta \rho < 0.2$ in some cases). Surprisingly neither phylogenetic signal nor sample size seemed to explain the strength of correlations observed between PD, SR and TD across traits. Phylogenetic signal also did not explain the degree to which PD was a superior predictor of TD than SR. The strength of observed correlations varied in a highly idiosyncratic way across traits in clade-level analyses, and further study is needed to investigate the factors that cause the diversity of a trait to show strong or weak correlations with SR and PD. In particular, factors such as the model of trait evolution that best describes variation in different traits or the degree to which various traits seem to be under stabilizing selection may be informative.

Geographical analyses showed that neither SR nor PD is correlated with adult body mass variance of regional assemblages. By contrast, both SR and PD were significantly correlated with geographical range variance (see the electronic supplementary material, table S2). In addition to differences in their biological and ecological significance, differences in data availability and the range of variation between range area and mass may also have affected their correlations with SR and PD.
Estimates of range area were available for almost all species of mammals, whereas estimates of adult body mass were available for only approximately 60 per cent of all species. In addition, range area spans 11 orders of magnitude in terrestrial mammals, whereas body mass spans only seven. We also found that within continental regions, correlations of PD and trait variance do frequently occur, but the relationship varies among different continents (see the electronic supplementary material, table S3). We suspect that these patterns are shaped at least partly by the evolutionary history of mammal assemblages that occur in each region. When we focused our analyses on areas where PD departs from what would be expected from SR, both PD and SR showed statistically significant correlations with TD (see the electronic supplementary material, table S4). In high-residual PD areas, mainly located in central Africa, PD is positively correlated with mass variance as expected from the clade-level analyses. Within these high-residual PD areas, species diverged from each other in the relatively distant past and so in many cases are more morphologically distinctive from each other than species occurring in other regions of the world [13,44]. Thus, as more species are sampled trait variance is likely to increase. By contrast, in areas such as western North America, Amazon basin and south Asian islands where PD is much lower than expected from SR, there is a negative correlation between PD and TD. This is probably because these areas contain clades that arose from recent speciation events, resulting in high redundancy of traits among species and leading to a reduction in trait variance as more (relatively similar) species are sampled [13,44].

Our analyses of geographical range variance showed very different results from those of body mass variance. Not only were consistent correlations among PD, SR and range area variance observed, but smaller scale analyses also showed that correlations between PD and geographical range variance were negative in areas of both high- and low-residual PD. This is likely related to the fact that areas of high- and low-residual PD are often also areas of high SR [13], and within these regions most species tend to have small geographical ranges (see the electronic supplementary material, figure S5, see also Gaston [45]). Thus, as more species are sampled species with similarly small geographical ranges will tend to be sampled, deflating variance. More generally, because geographical range area in mammals has very low phylogenetic signal (table 1), the correspondence between the two measures of TD (see GTA [5]; electronic supplementary material, figure S1), in cases where a measure of TD that is completely independent of SR is desired bin filling should not be used. We also note that an aspect of TD not directly quantified by either trait variance or bin filling is the range of species trait values (e.g. functional range sensu [34]). An ideal measure of niche filling would combine information on both the number and range of niches that are filled.

Finally, preliminary analyses showed that there is relatively little overlap between the geographical regions that show the highest levels of TD and PD, and relatively little correspondence between the two measures of TD (see the electronic supplementary material, figure S8). Current identified biodiversity hotspots also encompassed relatively few areas of high PD and trait variance. Somewhat surprisingly, however, they encompassed many areas where TD in terms of bin filling is highest, particularly in Asia. We speculate that such areas represent relatively intact assemblages where species with very small geographical ranges and with large body mass still occur. This represents an exciting area for future research.

In summary, our results support the idea that PD is often a reliable measure of the overall biodiversity of an assemblage, including TD. However, two caveats apply. First, PD will generally only be a useful surrogate measure of spatial patterns of trait variance in areas with higher PD than would be expected from their SR, and the difference in the variation explained by the PD versus SR alone is generally modest (e.g. PD only generally outperforms SR by $\Delta R^2 = 0.15–0.35$ in clade-level analyses). Second, PD and SR show similar correlations with bin filling. In cases where researchers are primarily interested in bin filling or similar measures of TD (e.g. niche filling), the additional effort needed to quantify PD compared with SR (i.e. estimating...
the phylogenetic relationships of species) may not be warranted. Nevertheless, PD remains the only measure of biodiversity that can simultaneously capture variation in SR, TD and genetic diversity above the species level.

We thank Nate Nibbelink, Jonathan Davies, Susanne Fritz and the Altizer laboratory group and Ezenwa lab group at Odum School of Ecology for comments. This work was partially supported by a NSF grant (no. 0642619) and the Odum School of Ecology.

REFERENCES


