bioRxiv preprint doi: https://doi.org/10.1101/060400; this version posted June 23, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

OPTIMAL INVENTORYING AND MONITORING OF TAXON, PHYLOGENETIC AND FUNCTIONAL DIVERSITY

Pedro Cardoso[1,2,*], José C. Carvalho[2,3], Luis C. Crespo[1,4] & Miquel A. Arnedo[4]

[1]*Finnish Museum of Natural History, University of Helsinki, P.O.Box 17 (Pohjoinen Rautatiekatu 13), 00014 Helsinki, Finland.*
[2]*CE3C – Centre for Ecology, Evolution and Environmental Changes / Azorean Biodiversity Group, Universidade dos Açores, 9700-042 Angra do Heroísmo, Portugal.*
[3]*Department of Biology, CBMA – Centre for Molecular and Environmental Biology, University of Minho, 4710-087 Braga, Portugal.*
[4]*Biodiversity Research Institute & Department of Animal Biology, Universitat de Barcelona, Av. Diagonal 645, 08028 Barcelona, Spain.*

* Correspondence, *E-mail: pedro.cardoso@helsinki.fi, Tel: (+358) 294128854.*

Abstract

1. Inventorying requires comprehensive sampling, while monitoring should build on it and focus on as few components as possible to detect changes in time or space. To sample or acquire information for both purposes can therefore be viewed as different processes, herewith called α-sampling and β-sampling. Biodiversity covers species, their genetic diversity and the way they interact in communities and ecosystems. Quantifying these aspects leads to the need to study taxon (TD), phylogenetic (PD) and functional diversity (FD). Here we propose and test a method for the optimization of sampling protocols to inventorying and monitoring TD, PD and FD for a wide range of taxa.

2. We used spiders as a case-study, although the same optimization procedure could be applied to any taxon requiring a combination of sampling methods. Spiders were sampled in multiple sites in the Iberian Peninsula. PD and FD were measured as the sum of the length of branches on a phylogenetic or functional tree. To reach the optimal combination of methods for inventory, we followed an iterative procedure optimizing the accumulation curve of diversity values. Optimal monitoring in space or time required minimizing the difference between sampled and true β-diversity. We used a framework that disentangles species replacement and richness differences contributions to overall β-diversity.

3. For α-diversity, the optimal combination of 24 samples for TD, PD and FD allowed sampling more than 50% of real diversity. For β-diversity, the optimal combination of six samples for all measures allowed reaching a bias (difference between sampled and true diversity) below 0.08.

4. The widespread use of similar sampling protocols worldwide and regular repetition in time can have a major impact on the scope and usefulness of global monitoring projects. In addition, the comparable sampling of biodiversity in space and/or time allows reusing of data collected for specific purposes, potentiating a synergistic effect among different projects. This makes data useful much beyond their initial plan. We strongly advocate the optimization, standardization and widespread adoption of sampling protocols for all taxa at a global level, for both inventory and monitoring of all levels of biodiversity: taxon, phylogenetic and functional.

bioRxiv preprint doi: https://doi.org/10.1101/060400; this version posted June 23, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Keywords: alpha diversity, Araneae, beta diversity, Iberian Peninsula, optimization, sampling, spiders.

Introduction

The current level of biodiversity loss is, so far, incommensurable, with our era being recognized as the grand stage to the sixth extinction (Leakey & Lewin 1996). Most probably more species go extinct than are described every single day. Besides a lack of insight and proactivity from politicians, stakeholders, scientists and the general public (Cardoso et al. 2011b), useful knowledge for the conservation of the vast majority of species is extremely scarce and the current population trends for most taxa are unknown (Cardoso et al. 2011a, b). This has vast implications much beyond the loss of taxa, as the ecosystem services they provide often are irreplaceable or require major funding to be recreated (Didham et al. 2016).

Inventorying and monitoring of biodiversity at all levels are some of the most urgent yet challenging tasks in ecology, biogeography and conservation. Although many researchers and practitioners tend to see these tasks as equivalent or interchangeable, the processes and goals are essentially different. If inventorying often requires comprehensive, time-consuming methods, monitoring should build on it and focus on as few components as possible that still allow the correct perception of changes on pre-determined indicators in time, or often, space (Table 1, Stork et al. 1996; Longino & Cowell 1997). That is to say, if inventorying requires assessing alpha diversity on a given, usually small, site and time-frame, monitoring requires assessing beta diversity, i.e., changes in species composition across sites and/or over time. To sample or acquire useful information for both purposes can therefore be viewed as fundamentally different processes, herewith called α and β-sampling (Table 1). Establishing a parallel with taxonomy, α-sampling is equivalent to a species description, requiring maximum detail, while β-sampling is similar to a diagnosis, where only characteristics relevant for comparison with other, often look-alike, species are emphasized.

Table 1. Characteristics of α and β-sampling.

	α-sampling	**β**-sampling
Purpose	inventorying	monitoring
Target	maximize sampled α-diversity	minimize difference between sampled and true β-diversity
Taxonomic scope	All taxa	Selected taxa
Sampling effort	Major (>= 50% completeness)	Minor (often < 50% completeness)
Geographic scope	Global	Selected sites
Seasonality	Full year	Selected season (often of maximum diversity)
Periodicity	Single or sporadic (> 10 years)	Regular (1 to 10 years)

On the first place, both inventorying and monitoring of biodiversity require defining what biodiversity is. Its usual meaning covers the variety of species, their genetic variation and interrelation and the way they interact to constitute complex assemblages of communities and ecosystems. Quantifying all these aspects leads to the need to study three different facets of biodiversity: taxon, phylogenetic and functional diversity (covering species, genes and ecosystems, respectively).

bioRxiv preprint doi: https://doi.org/10.1101/060400; this version posted June 23, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Taxon Diversity (TD) is quantified based on the number of taxa and, often, on the distribution of abundances. TD treats taxa as being equally distinct from one another and disregards the fact that communities are composed of species with different evolutionary histories (Webb et al. 2002; Graham & Fine 2008) and a diverse array of ecological functions (Cardoso et al. 2014a).

Phylogenetic Diversity (PD) takes the evolutionary relationships between taxa into account and reflects how much evolutionary history is behind the species in the community. Communities with identical TD may differ widely with respect to their evolutionary past, depending on the time of divergence of species from their nearest common ancestor (Webb et al. 2002; Graham & Fine 2008). PD has mostly been measured based on phylogenetic trees or cladograms (Faith 1992). This notion can be extended to intra-specific variability, in which case individuals or populations of one or more species are studied under the same framework (Ståhls et al. 2016). Yet, in lieu of precise phylogenetic information on the member taxa of a community, taxonomic diversity (LD) may be used as a surrogate for PD, being calculated from an ultrametric tree (i.e. a dissimilarity matrix) representing the Linnaean hierarchy (Ricotta et al. 2012).

Functional diversity (FD) quantifies the components of biodiversity that influence how an ecosystem operates or functions (Tilman et al. 2001) and reflects the amount of biological functions or traits displayed by the species occurring in given communities. Communities with completely different species composition may be characterized by low variation in functional traits, with unrelated species replacing others with similar roles in the network (Villéger et al. 2012). FD has also been quantified in many different ways (e.g. Rao 1982), often using approaches that were first developed for PD (Petchey & Gaston, 2002, 2006). Using tree-based approaches to quantify FD does however present a major advantage, it allows the direct comparison between TD, PD and FD under the same framework (Cardoso et al. 2014a, b, 2015).

A number of global initiatives are currently underway for the inventory or monitoring of biodiversity. Monitoring has however been the main concern, given the current rate of global change. Some of the heavy users of monitoring data are the GEO BON - Group of Earth Observation Biodiversity Observation Network, within which Pereira et al. (2010) proposed a global terrestrial species monitoring program, advocating that vascular plants and birds should be the candidate taxa for that purpose. The GEO BON is now expanding to include other taxa such as butterflies (van Swaay et al. 2015). The IUCN - International Union for the Conservation of Nature, proposed the Red List Index (RLI), applicable to any organism group and already calculated for birds (Butchart et al. 2004; Hoffman et al. 2010), mammals (Hoffmann et al. 2010, 2011), amphibians (Hoffman et al. 2010), corals (Butchart et al. 2010) and cycads (Anonymous 2015). Finally, the Living Planet Index (LPI) is based on population trends of vertebrates only (Loh et al. 2005; Collen et al. 2009).

All these global initiatives have a number of points in common. Firstly, they do not currently consider the majority of taxa, and there is no evidence that the ones covered can serve as good surrogates for detecting trends in the remaining of biodiversity. In fact, Pereira et al. (2013) noted that Essential Biodiversity Variables relating to species, genes or traits require representative sampling across taxonomic groups, recognizing the difficulty in obtaining such data. Second, these initiatives often do not explicitly consider phylogenetic and/or functional diversity. The lack of knowledge on most species is a major hurdle and probably the main reason for the limitations of current initiatives. The existence of optimized protocols to inventory and monitor TD, PD and FD of a much wider range of

bioRxiv preprint doi: https://doi.org/10.1101/060400; this version posted June 23, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

taxa is therefore essential if we really want to follow the current trends in biodiversity as a whole, from local to global scales (Didham et al. 2010; Cardoso et al. 2011b).

Cardoso (2009) proposed a method to optimize the inventory of species (TD), particularly focusing on mega-diverse groups. This former study demonstrated that it is possible to sample in a standardized, yet optimized, way, and the protocol proposed just 7 years ago is now being used worldwide to sample spiders (Cardoso et al. 2009b; Carvalho et al. 2011a, b, 2012a, b; Crespo et al. 2009, 2013; Emerson et al. subm.; Arnedo et al. in prep; Malumbres-Olarte et al. in prep.). Here we extend the concept from inventorying to monitoring and its application from TD to PD and FD.

Materials and methods

Case-study

In the present work, we use spiders as a case-study, although the same optimization procedure could be applied to any taxon that requires a combination of sampling methods to be fully represented, from plants to mammals. Spiders are among the most numerous arthropods of many samples in all kinds of habitats. More than 46 000 species of spiders are known (World Spider Catalogue 2016), which makes them one of the most diverse orders worldwide, ranking 7th after the largest insect orders and Acari, the ticks and mites. Besides global diversity, local diversity can also be extremely high, with several hundreds of species co-existing in areas as small as 1 ha, especially in high-diversity biomes such as tropical forests (Coddington et al. 2009).

Our team have sampled Iberian spiders in 29 sites using a common, standardized and optimized, protocol since 2004 (Table 2). In two forest sites, Arrábida and Gerês, sampled in 2004 and 2005 respectively, we collected a much larger number of samples (320) than in all others and these were used previously to define the current inventory protocol for TD (Cardoso et al. 2008a, b; Cardoso 2009). The remaining 27 sites were sampled between 2006 and 2014 mostly using a “low effort” sub-protocol (24 samples) intended to capture 50% of the estimated species richness (Cardoso et al. 2009b; Crespo et al. 2009, 2013; Carvalho et al. 2011a, b, 2012a, b; Arnedo et al. in prep.). Numerous habitat types were sampled using a combination of methods: aerial searching, beating trees and branches, ground searching, sweeping and pitfall trapping (for details of the sampling methods see Cardoso et al. 2008a, b, 2009c). All samples (except pitfall traps) comprised one-hour of active sampling, measured with a stopwatch and were performed both day and night. Diurnal and nocturnal samples of each method were considered different methods as assemblages assessed are considerably different. Activity not directly involved in sampling was excluded by pausing the watch (e.g., travel time to a different area within the subplot, logistical problems, equipment maintenance, personal tasks). Pitfall traps were pooled in groups of four as we estimated that four traps take an average of one hour to set up (although this is extremely variable, from 10 minutes in sandy dunes to 2 hours in rocky ground) and collect after two weeks, so that they were comparable with other methods in terms of sampling effort. Sampling always occurred between May and July, when spider species and abundance typically are highest in the Mediterranean (Cardoso et al. 2007).

bioRxiv preprint doi: https://doi.org/10.1101/060400; this version posted June 23, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Table 2. 29 sampled sites and respective characteristics. At all sites at least 24 samples were performed. In forest sites sampling always included 12 pitfall trap samples (equivalent to 48 traps), 4 aerial samples during the night and 2 beating and sweeping samples during day and night each. In other habitats with no tree cover aerial sampling was replaced by ground sampling and beating samples by extra sweeping samples.

Site	Latitude	Longitude	Habitat	Elevation (m)	Reference
Aiguestortes 1	42.5496	0.8725	Oak forest	1760	Arnedo et al. in prep.
Aiguestortes 2	42.5491	0.8714	Oak forest	1740	Arnedo et al. in prep.
Amorosa	41.6566	-8.8239	Dunes	10	Carvalho et al. 2011a, b, 2012a, b
Arrábida	38.5067	-8.9752	Oak forest	60	Cardoso et al. 2008a
Cabañeros 1	39.3566	-4.3591	Oak forest	760	Arnedo et al. in prep.
Cabañeros 2	39.3516	-4.3589	Oak forest	740	Arnedo et al. in prep.
Cabañeros 3	39.3618	-4.4173	Oak forest	770	Arnedo et al. in prep.
Cabañeros 4	39.3634	-4.4170	Oak forest	770	Arnedo et al. in prep.
Esposende	41.5492	-8.7900	Dunes	10	Carvalho et al. 2011a, b, 2012a, b
Gerês	41.7953	-8.1363	Oak forest	660	Cardoso et al. 2008b
Guadiana	37.5750	-7.5350	Shrubland	120	Cardoso et al. 2009c
Monchique 1	37.3196	-8.5975	Shrubland	850	Crespo et al. in prep.
Monchique 2	37.3250	-8.5552	Oak forest	540	Crespo et al. in prep.
Monfragüe 1	39.8330	-6.0641	Oak forest	320	Arnedo et al. in prep.
Monfragüe 2	39.8280	-6.0325	Oak forest	320	Arnedo et al. in prep.
Ordesa 1	42.6068	0.1313	Oak forest	1400	Arnedo et al. in prep.
Ordesa 2	42.5943	0.1529	Oak forest	1160	Arnedo et al. in prep.
Paúl de Arzila 1	40.1792	-8.5540	Shrubland	10	Crespo et al. 2009
Paúl de Arzila 2	40.1746	-8.5552	Riparian forest	10	Crespo et al. 2009
Picos de Europa 1	43.1445	-4.9267	Oak forest	1070	Arnedo et al. in prep.
Picos de Europa 2	43.1777	-4.9058	Oak forest	760	Arnedo et al. in prep.
Picos de Europa 3	43.1435	-4.9488	Oak forest	1100	Arnedo et al. in prep.
Picos de Europa 4	43.1723	-4.9086	Oak forest	940	Arnedo et al. in prep.
Santo André 1	38.0267	-8.8336	Dunes	10	Carvalho et al. 2011a, b, 2012a, b
Santo André 2	38.0810	-8.7809	Oak forest	20	Crespo et al. in prep.
São Mamede 1	39.4194	-7.3388	Shrubland	520	Crespo et al. in prep.
São Mamede 2	39.2936	-7.3836	Oak forest	620	Crespo et al. in prep.
Sierra Nevada 1	36.9615	-3.4188	Oak forest	1790	Arnedo et al. in prep.
Sierra Nevada 2	37.1838	-3.2628	Oak forest	1710	Arnedo et al. in prep.

Phylogenetic and Functional Diversity

PD and FD were measured as the sum of the length of branches on a phylogenetic or functional tree respectively (Faith 1992; Petchey & Gaston 2002). Since TD can also be represented by a tree with each taxon linked directly to the root by a branch of unit length (star tree), tree diagrams provide a common basis for sampling optimization of TD, PD and

bioRxiv preprint doi: https://doi.org/10.1101/060400; this version posted June 23, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FD (Cardoso et al. 2014a, b, 2015).

As no phylogenetic tree was available for all or even a majority of species, we used the taxonomic hierarchy as a surrogate, with species, genera, families and sub-orders separated by 0.25, 0.5, 0.75 and 1 unit respectively (see also Ricotta et al. 2012). For FD we relied on body size (average between males and females), type of web (sensing, tube, sheet, orb or no web), type of hunting (active or ambushing), prey range (stenophagous or euryphagous), vertical stratification (ground or vegetation) and daily activity (diurnal or nocturnal). All traits were collected from literature (see also Cardoso et al. 2011c) or derived from our own sampling data. To build a functional dissimilarity matrix between species we used Gower's distance which allows mixing continuous and categorical data (Pavoine et al. 2009). The functional tree was built using UPGMA as in Cardoso et al. (2011c). All analyses were performed using the R language (R Core Team 2016).

The general optimization algorithm

To reach the optimal combination of methods we followed an iterative procedure that minimized the bias of sampled diversity (TD, PD or FD) for any undersampling level compared with true diversity as revealed by full sampling (equivalent to maximizing richness in the case of α-sampling, see example in Fig. 1):

1. Start with 0 samples (an empty pool);
2. Add one randomly chosen sample from a given method (without replacement);
3. Calculate the absolute difference between sampled and true diversity (i.e. bias);
4. Repeat from step 2 many times (e.g., 10 000, we recommend at least 1000);
5. Calculate the average bias if adding one sample of the method being tested;
6. Repeat steps 2–5 for each method in turn;
7. Identify which method, on average, minimizes bias;
8. Add 1 sample of such method to the pool;
9. Repeat from step 2 until available samples are exhausted.

Optimization of inventory protocols (α-sampling)

Only the two thoroughly sampled forest sites were used to guarantee a close to complete inventory in the analyses. The original algorithm (Cardoso 2009) maximized the species richness (TD) per given total number of samples at each site individually. Here we first adapted it to allow maximizing PD and FD in addition to TD. Moreover, we optimized multiple sites simultaneously by using the average proportion of species sampled across sites as the measure of efficiency. This way richer and poorer sites were given the same importance in determining the optimal combination.

bioRxiv preprint doi: https://doi.org/10.1101/060400; this version posted June 23, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

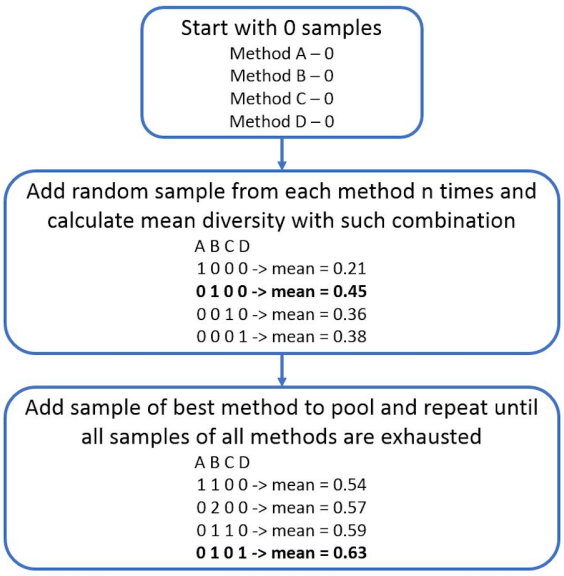

Fig. 1. Scheme of the optimization algorithm. Best method to add at each step for both α and β diversity is the one which minimizes bias, i.e., the difference between sampled and true diversity. In the case of α, this is equivalent to maximizing sampled diversity (as shown in figure).

To evaluate the optimality of the TD, PD and FD α-sampling protocols we first plotted the accumulation curves of diversity as a function of the number of samples (up to 320) and compared them with 10 000 randomly chosen sets of samples. The accumulation of samples per method was also graphically checked. Then we compared the different protocols at 24 samples, as this number was shown to sample around 50% of TD (Cardoso 2009) and we wanted to verify if the optimal combination would guarantee a similar or higher proportion of sampled PD and FD.

Finally, we wanted to simulate building a protocol based on a predetermined set of samples. This may be useful when we want to improve on an existing sampling protocol, yet guaranteeing a minimum common denominator to allow comparisons between old and new. This implied forcing the optimization exercise to start with such common denominator (we chose 12 pitfall trap samples as an example) and searching for the methods that would better complement it up to a total of 24 samples. We then checked if the final solution was similar to the unconstrained optimization. For this exercise we only tested TD. All analyses were performed using the R package BAT (Cardoso et al. 2015).

Optimization of monitoring protocols (β-sampling)

As mentioned previously, optimal inventorying requires the maximization of sampled α-diversity for a given sampling effort (Table 1). This is equivalent to minimizing the difference between sampled and true α-diversity. Conversely, optimal monitoring in space or time requires minimizing the difference between sampled and true β-diversity. This new algorithm allows maximizing the chances of detecting changes across sites or over time for any taxon and in any biome, using different collecting methods, as long as the effort per sample is comparable, while minimizing the sampling costs and resources employed. To

bioRxiv preprint doi: https://doi.org/10.1101/060400; this version posted June 23, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

guarantee such goal one can use any β-diversity measure of interest. Yet, here we used a recently developed framework that disentangles the two antithetic processes that drive overall β-diversity: species replacement and species richness differences (Carvalho et al. 2012b). This framework was further expanded to PD and FD (Cardoso et al. 2014a) and can be represented as:

$$\beta_{total} = \beta_{repl} + \beta_{rich}$$

where β_{total} represents the total community variation, β_{repl} gives its fraction resulting from species replacement, i.e., the substitution of species across sites or over time, and β_{rich} accounts for the variation as a result of species richness differences, determined by the net loss / gain of species across sites or over time. These measures can be applied site-by-site (pairwise approach) or using multiple-site versions based on presence-absence or abundance data (see Legendre 2014).

This framework is ecologically meaningful and mathematically correct, as it satisfies three important criteria (Carvalho et al. 2012b, 2013): (1) the indices are additive because they are scaled in relation to the same denominator; (2) the β_{repl} and β_{rich} metrics present a monotonic behaviour in relation to increases in the amounts of replacement and richness differences in the data; (3) the relative magnitude of β_{repl} and β_{rich} metrics is proportional to the relative amounts of replacement and richness difference in the data.

Monitoring programs are seldom carried out on continuous landscapes. Samples are usually collected from different sized patches. It is well known that larger patches contain more species (species-area relationship), therefore changes in species composition may arise solely due to richness differences among patches. Moreover, temporal changes in species richness (alpha diversity) or replacement of some species with others (e.g. natives by exotics that outcompete them) are also likely in long-term monitoring programs. In this context, the beta diversity framework here proposed is particularly suitable since allows disentangling dissimilarity due to species replacement from dissimilarity due to species richness differences determined by area effects or species loss (or gain) over time.

These measures of beta diversity are also particularly robust to biases due to undersampling (Cardoso et al. 2009a) and therefore we expected the sampling effort necessary to attain low bias to be small compared with α-sampling. It is also a necessary property of any monitoring protocol to be easily replicated in time, so the overall number of samples could never be as high as for inventorying (24 samples). We could then use all pairwise combinations among the 29 sites inventoried, as they all had 24 samples as a minimum.

The target for the algorithm was to minimize the summed absolute difference between sampled and true β-diversity values using all three measures for TD, PD and FD. Sampled diversity was calculated at all levels of undersampling, from 1 to 23 samples. As true diversity we used the values attained at 24 samples. Given that six is the maximum number of samples most collectors can perform daily without tiring to the point of biasing the collection, we quantified the bias in β-diversity expected if only six samples were made per site. Otherwise we made similar analyses as for α-sampling, namely the optimization of TD, PD and FD using multiple sites simultaneously, plotted the accumulation curves of diversity as a function of number of samples (up to 24) and compared them with the average accumulation of 1000 randomly chosen sets of samples. The accumulation of samples per method was also graphically checked. Then we compared the different

bioRxiv preprint doi: https://doi.org/10.1101/060400; this version posted June 23, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

protocols at six samples and finally simulated the case of a protocol building upon a predetermined set of samples (this time starting with four pitfall trap samples). All analyses were performed using the R package BAT (Cardoso et al. 2015).

Results

A total of 272 species and 13064 adult individuals (from two sites) were used for α-diversity and 587 species and 18623 adults (from 29 sites) for β-diversity analyses. This was the largest combined effort ever made to sample spiders in a standardized way in the Iberian Peninsula.

Table 3. Optimal number of samples per method for inventorying (α-sampling) and monitoring (β-sampling) of spider taxon (TD), phylogenetic (PD) and functional (FD) diversity. A total of 24 and six samples were chosen respectively (see text for details). The constrained runs start with 12 or four pitfall trap samples respectively.

		Aerial		Beating		Ground		Pitfall	Sweep		Average proportion of diversity (**α**) or bias (**β**)
		Day	Night	Day	Night	Day	Night		Day	Night	
α-sampling	TD	0	3	3	1	0	0	10	0	7	0.507
	PD	0	5	4	0	0	0	8	0	7	0.571
	FD	0	7	1	0	5	0	9	0	2	0.782
	TD constrained	0	2	3	1	0	0	12	1	5	0.502
β-sampling	TD	-	1	1	0	-	-	2	2	0	0.060
	PD	-	0	2	0	-	-	3	1	0	0.062
	FD	-	0	2	0	-	-	2	2	0	0.079
	TD constrained	-	0	1	0	-	-	4	1	0	0.064

For α-diversity, the optimal combination of 24 samples for both TD and PD was very similar (Table 3) and close to the one already found for TD in a former study (Cardoso 2009). In both cases the proportion of diversity sampled was higher than 50%, reaching 57% of PD. The optimal protocol was close or even over the upper confidence limit for random sampling for most of the accumulation process (Fig. 2). For FD the main difference laid on the optimality of doing ground searching during the day instead of beating or sweeping. In any case, as many species are very similar in their traits, the sampled diversity with just 24 samples was very high, reaching 78%. This was confirmed by the much faster accumulation rate of FD with increasing number of samples (Fig. 2). In the constrained exercise, the algorithm was able to rapidly reach method combinations and sampling levels on par with the unconstrained optimization (Table 3; Fig. 2).

bioRxiv preprint doi: https://doi.org/10.1101/060400; this version posted June 23, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

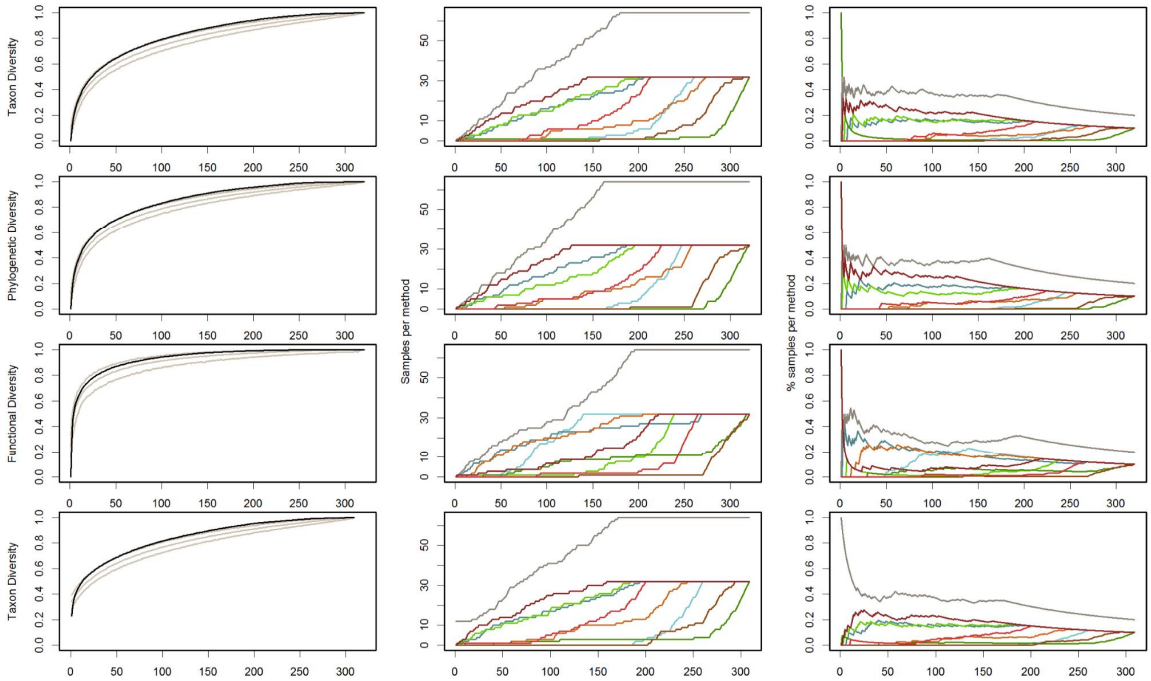

Fig. 2. Accumulation curves for optimized inventorying (α-sampling) of taxon, phylogenetic and functional diversity. Values compared with random number of samples per method (left panels, grey lines are the median and 95% confidence limits), number of samples per method (center panels), and proportion of samples per method (right panels). Last row is for taxon diversity constrained to 12 pitfall trap samples. Methods: aerial day (light blue), aerial night (dark blue), beating day (light green), beating night (dark green), ground day (light brown), ground night (dark brown), pitfall traps (grey), sweeping day (light red), sweeping night (dark red).

For β-diversity, the optimal combination of six samples for all measures was similar (Table 3), including two beating and two sweeping samples during the day and two pitfall trap samples (corresponding to eight traps left in the field for two weeks). In all cases the bias from true diversity was between 0.060 (for TD) and 0.079 (for FD). The optimal protocol was however not significantly different from a random one based on the inventory protocol (Fig. 3). In the constrained exercise, the algorithm was able to rapidly reach method combinations and sampling levels on par with the unconstrained optimization (Table 3; Fig. 3). It should be noted that in cases of open habitat, with no or little tree cover, the beating samples should be replaced by two extra sweeping samples to keep a similar total number of samples (as in Cardoso 2009).

bioRxiv preprint doi: https://doi.org/10.1101/060400; this version posted June 23, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

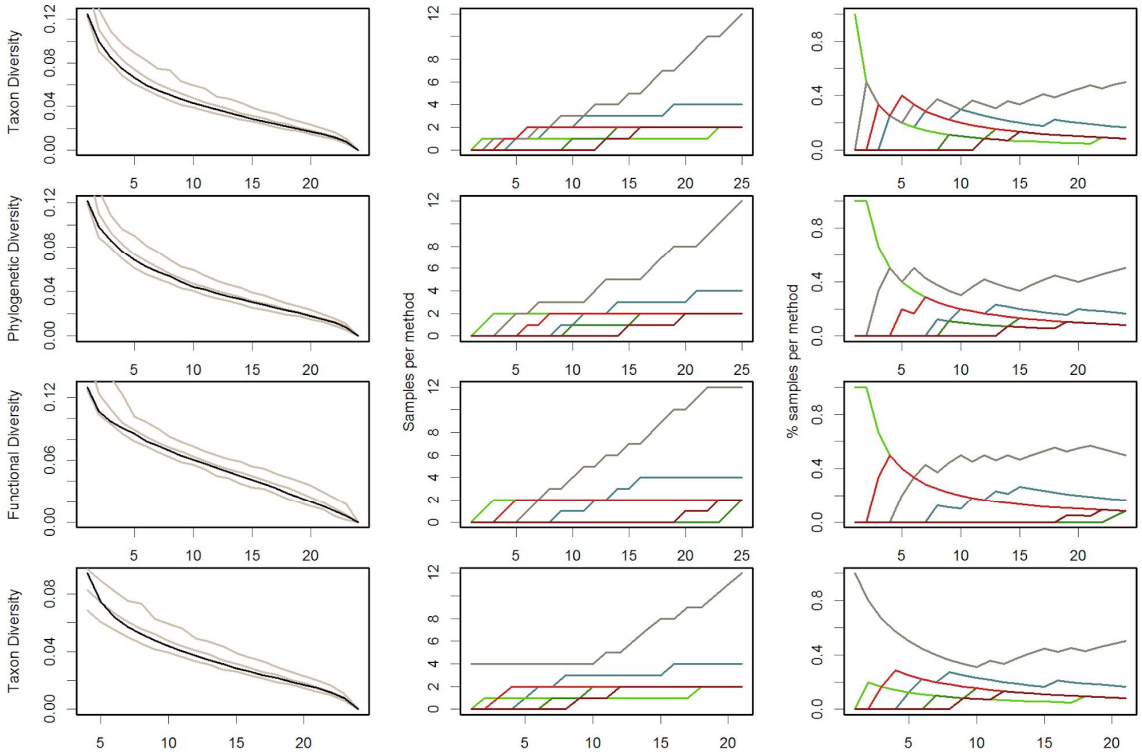

Fig. 3. Accumulation curves for optimized monitoring (β-sampling) of taxon, phylogenetic and functional diversity. X-axis of left graphics is the average bias from true beta-diversity among 1000 runs. Last row is for taxon diversity constrained to four pitfall trap samples. Other values and line colors as in Fig. 2.

Discussion

Here we propose a method to optimize and standardize the inventory and monitoring of taxon, phylogenetic and functional diversity, and further introduce the definitions of α-sampling and β-sampling. We demonstrate the method using Iberian spiders as a case-study. This same method is applicable to any taxon at any region, as long as sampling can be done using a combination of methods that are comparable in effort.

Instead of relying on subjective opinions of researchers, this approach allows the proposed protocols to meet six requirements that facilitate their widespread use. The protocols are suitable for the purpose and target organisms, they are efficient by maximizing the gain/effort ratio, feasible by requiring minimum resources, flexible as different sampling efforts can be tested and proposed, transparent as the process to reach them is explicit and accountable as results were and can be tested a posteriori (Cardoso 2009).

Inventorying or monitoring?

The answer to this question is most of the times obvious. An inventory is required if the goal is to know which species, clades or traits live where, to study community assemblage

bioRxiv preprint doi: https://doi.org/10.1101/060400; this version posted June 23, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

patterns and drivers or to prioritize areas for conservation. As the aim is to maximize sampled α-diversity, all taxa of interest should be targeted, a high level of sampling completeness is needed and sampling should cover as much as possible the entire region and seasons (Table 1). On the other hand, this is either a one-time or a rare event, being almost invariably impossible to replicate in the short or even medium-term.

In contrast, monitoring is required if the goal is to detect and if possible quantify changes in community composition (or of a single species of particular interest) in space or, more often, over time (Noss 1990; Stork et al. 1996; Bawa & Menon 1997; Hilty & Merenlender 2000; Yoccoz et al. 2001; Danielsen et al. 2005). As the aim is to minimize difference between sampled and true β-diversity, one may focus on selected taxa, completeness is not a requirement, and sampling can cover only a subset of sites sampled in the height of seasonal abundance for the selected taxa (Table 1). This minimizes the required time and resources and allows the frequent replication in time.

The monitoring protocol we propose was delineated using sites from contrasting habitats (Table 2). This probably explains why a combination of several different methods is optimal, as different components of the assemblages should be targeted. If a single habitat type was used, it is possible that a simplified approach, using less methods and samples per method, was needed. And each habitat type would require a different combination of methods. Optimizing β-sampling may therefore depend on the goals and should be done for each biome, region and/or habitat independently. Contrary to inventorying, which can be, and is being done, at a global level using very similar methodology, allowing comparisons at large to global scales, monitoring is case-specific, and results are context-dependent.

Nevertheless, ideally, before starting any kind of monitoring of a site or region an inventory should be done, so that a baseline can be established (Stork et al. 1996). Inventories provide the baseline that allows perceiving if changes in community composition are due to changes in the abundances of species, clades or traits already living but previously rare in the site, or to the replacement (i.e. β_{repl}), extirpation or arrival (i.e. β_{rich}) of new components to the community.

Uncoupling of real vs sampling artefact related change in monitoring protocols is a hard task and deserves careful consideration. First, the data used to define the protocol is of critical importance. It has to be extensive, from multiple sites covering the entire spectrum we want to monitor, to provide as good baseline as possible. Second, we must be careful choosing the β-diversity measures used to optimize the protocol. The ones used here are much less prone to error due to undersampling than either α-diversity or other β-diversity measures (Cardoso et al. 2009a).

Finally, one may combine simultaneous α and β-sampling. One may α-sample a number of reference sites and β-sample additional sites with the sole purpose of comparing them with the reference condition. This might be useful for example to test change in community composition along a disturbance gradient, test the effect of habitat margins, or study meta-population source-sink dynamics in a fragmented landscape.

Quantifying phylogenetic and functional diversity

We chose to focus on trees or dendrograms as representations for both PD and FD, not only because trees allow a straightforward adaptation of existing algorithms for TD, but also because they allow the comparison of TD, PD and FD under a common framework (Cardoso et al. 2014a, b, 2015). Although there is no other obvious alternative for PD, this

bioRxiv preprint doi: https://doi.org/10.1101/060400; this version posted June 23, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

optimization method can be adapted to other representations such as multidimensional hypervolumes, a common representation for FD. Yet, due to the nature of the calculation of convex-hull volumes, each sample may occupy a relatively small portion of the functional space, yet when considering several samples simultaneously, the space in between samples may be occupied, even when none of the samples individually occupies it. To what extent this shortcoming causes biases in practice remains to be studied. Additionally, the calculation of the convex-hull volume for each sample requires that the number of species always exceeds the number of traits (Laliberté & Legendre 2010), which constrains the user to reduce the dimensionality of the functional space, as many single samples will often have few species and yet must be included in analyses.

Setting up a sampling scheme

A number of considerations must be made before the replication of any sampling protocol. The spatial extent and temporal scale must be determined in view of specific objectives and available resources. The number and location of sampling sites could, as much as possible, be consistent with some procedure common for an entire scheme. Many projects may benefit from maximizing environmental diversity covered by sampling (including habitat types), with the reasoning that environmental and biological diversity may be correlated (Faith & Walker 1996; Hortal & Lobo 2005). This allows maximizing the overall diversity captured and, most importantly, the range of variation in biodiversity of a region. One must remember though, that if monitoring is to be done, the sites should be fixed and ease of accessibility may be important.

For monitoring, the frequency and duration of the scheme is crucial. Often it is not possible to detect changes in species composition or population abundance before a relatively large time lag. It is also important to be able to tell apart consistent change and natural fluctuations. Thus, besides specimen sampling, it is important to collect the most relevant biotic and abiotic variables. Usually habitat type, temperature, precipitation, tree and ground cover during or immediately before the time of sampling could potentially be useful to help understanding community composition and change. Precipitation in particular is known to strongly influence the activity of most species and collectors should therefore make an effort to only sample in appropriate conditions.

Voucher specimens, or better yet entire collections, should be preserved as only this way it is possible to guarantee future comparability of projects if taxonomy changes or if further analyses are needed relying on genetic or functional data. One may for example want to sequence new genes that allow the comparison of intra-specific genetic diversity of given species between sites or years even if these comparisons were not thought of from the beginning of the project. Or to detect changes in functional traits such as body size driven by environmental change. All projects involving specimen sampling should therefore plan and allocate appropriate funding to specimen deposition in public collections such as in natural history museums (Schilthuizen et al. 2015).

Going global

The widespread use of similar sampling protocols at global levels and regular repetition in time can potentially have a major impact on the scope and consequently usefulness of projects such as GEO-BON, RLI and LPI. But the impact goes much beyond these. Often

bioRxiv preprint doi: https://doi.org/10.1101/060400; this version posted June 23, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

researchers use sampling methods and effort per method with which they are more familiar with or that are specific to a given project, necessarily limited in space and time. Besides potentially low cost-effectiveness, this creates data that are not comparable with other datasets collected by different or even the same researchers for other projects. The comparable sampling of biodiversity in space and/or time allows reusing of data collected for specific purposes, potentiating a synergistic effect among different projects. For spiders, our team is now working on multiple analyses that use data collected during the last decade by different teams for different purposes. Ultimately, all these studies will be comparable at a large scale. Similar tools are being developed for savannas (Muelelwa et al. 2010), tropical forests (Malumbres-Olarte et al. subm.) and oceanic islands (Borges et al. in prep., Emerson et al. in prep.). This makes data useful much beyond their initial plan. We therefore strongly advocate the optimization, standardization and widespread adoption of sampling protocols for all taxa at a global level, for both inventory and monitoring of all levels of biodiversity: taxon, phylogenetic and functional.

Acknowledgments

We thank all the collectors of past and present studies that contributed to the dataset used in this paper: Rui Carvalho, Alberto de Castro, Cristina Frías-López, Pedro Humberto Castro, Clara Gaspar, Ana Filipa Gouveia, Sérgio Henriques, Nuria Macías-Hernández, Eva de Mas, Paola Mazzuca, Elisa Mora, Jordi Moya, Vera Opatova, Luis Carlos Pereira, Enric Planas, Carles Ribera, Marcos Roca-Cusachs, Cristina Rufino, Dolores Ruiz, Nikolaj Scharff, Jesper Schmidt, Israel Silva, Ricardo Silva, Pedro Sousa, Tamás Szüts, Vanina Tonzo and Sandra Videira. JCC was supported by the Luso-American Foundation (FLAD C&T Links 2010 grant). MA was supported by grant 485/2012 of the Organismo Autónomo de Parques Nacionales (Ministerio de Agricultura, Alimentación y Medio Ambiente).

References

Anonymous (2015) The Millennium Development Goals Report. United Nations, New York.

Bawa, K.S. & Menon, S. (1997) Biodiversity monitoring: the missing ingredients. Trends in Ecology and Evolution, 12: 42-43.

Butchart, S.H.M., Stattersfield, A.J., Bennun, L.A., Shutes, S.M., Akcakaya, H.R., Baillie, J.E.M., Stuart, S.N., Hilton-Taylor, C. & Mace, G.M. (2004) Measuring global trends in the status of biodiversity: Red List Indices for birds. PLoS Biology, 2: e383.

Butchart, S.H.M., Walpole, M., Collen, B., van Strien, A., Scharlemann, J.P.W., Almond, R.E.A., Baillie, J.E.M., Bomhard, B., Brown, C., Bruno, J., Carpenter, K.E., Carr, G.M., Chanson, J., Chenery, A., Csirke, J., Davidson, N.C., Dentener, F., Foster, M., Galli, A., Galloway, J.N., Genovesi, P., Gregory, R.D., Hockings, M., Kapos, V., Lamarque, J.-F., Leverington, F., Loh, J., McGeoch, M., McRae, L., Minasyan, A., Morcillo, M.H., Oldfield, T.E.E., Pauly, D., Quader, S., Revenga, C., Sauer, J.R., Skolnik, B., Spear, D., Stanwell-Smith, D., Stuart, S.N., Symes, A., Tierney, M., Tyrrell, T.D., Vié, J.-C. & Watson, R. (2010) Global biodiversity: indicators of recent

bioRxiv preprint doi: https://doi.org/10.1101/060400; this version posted June 23, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

declines. Science, 328: 1164-1168.

Cardoso, P. (2009) Standardization and optimization of arthropod inventories – the case of Iberian spiders. Biodiversity and Conservation, 18: 3949-3962.

Cardoso, P., Silva, I., Oliveira, N.G. & Serrano, A.R.M. (2007) Seasonality of spiders (Araneae) in Mediterranean ecosystems and its implications in the optimum sampling period. Ecological Entomology, 32: 516-526.

Cardoso, P., Gaspar, C., Pereira, L.C., Silva, I., Henriques, S.S., Silva, R.R. & Sousa, P. (2008a) Assessing spider species richness and composition in Mediterranean cork oak forests. Acta Oecologica, 33: 114-127.

Cardoso, P., Scharff, N., Gaspar, C., Henriques, S.S., Carvalho, R., Castro, P.H., Schmidt, J.B., Silva, I., Szüts, T., Castro, A. & Crespo, L.C. (2008b) Rapid biodiversity assessment of spiders (Araneae) using semi-quantitative sampling: a case study in a Mediterranean forest. Insect Conservation and Diversity, 1: 71-84.

Cardoso, P., Borges, P.A.V. & Veech, J.A. (2009a) Testing the performance of beta diversity measures based on incidence data: the robustness to undersampling. Diversity and Distributions, 15: 1081-1090.

Cardoso, P., Crespo, L.C., Carvalho, R., Rufino, A.C. & Henriques, S.S. (2009b) Ad-hoc vs. standardized and optimized arthropod diversity sampling. Diversity, 1: 36-51.

Cardoso, P., Henriques, S.S., Gaspar, C., Crespo, L.C., Carvalho, R., Schmidt, J.B., Sousa, P. & Szűts, T. (2009c) Species richness and composition assessment of spiders in a Mediterranean scrubland. Journal of Insect Conservation, 13: 45-55.

Cardoso, P., Borges, P.A.V., Triantis, K.A., Ferrández, M.A. & Martín, J.L. (2011a) Adapting the IUCN Red List criteria for invertebrates. Biological Conservation, 144: 2432-2440.

Cardoso, P., Erwin, T.L., Borges, P.A.V. & New, T.R. (2011b) The seven impediments in invertebrate conservation and how to overcome them. Biological Conservation, 144: 2647-2655.

Cardoso, P., Pekár, S., Jocqué, R. & Coddington, J.A. (2011c) Global patterns of guild composition and functional diversity of spiders. PLoS One, 6: e21710.

Cardoso, P., Rigal, F., Carvalho, J.C., Fortelius, M., Borges, P.A.V., Podani, J. & Schmera, D. (2014a) Partitioning taxon, phylogenetic and functional beta diversity into replacement and richness difference components. Journal of Biogeography, 41: 749-761.

Cardoso, P., Rigal, F., Borges, P.A.V. & Carvalho, J.C. (2014b) A new frontier in biodiversity inventory: a proposal for estimators of phylogenetic and functional diversity. Methods in Ecology and Evolution, 5: 452-461.

Cardoso, P., Rigal, F. & Carvalho, J.C. (2015) BAT - Biodiversity Assessment Tools, an R package for the measurement and estimation of alpha and beta taxon, phylogenetic and functional diversity. Methods in Ecology and Evolution, 6: 232-236.

Carvalho, J.C., Cardoso, P., Crespo, L.C., Henriques, S., Carvalho, R. & Gomes, P. (2011a) Biogeographic patterns of spiders in coastal dunes along a gradient of mediterraneity. Biodiversity and Conservation, 20: 873-894.

Carvalho, J.C., Cardoso, P., Crespo, L.C., Henriques, S., Carvalho, R. & Gomes, P. (2011b) Determinants of beta diversity of spiders in coastal dunes along a gradient of mediterraneity. Diversity and Distributions, 17: 225-234.

Carvalho, J.C., Cardoso, P., Crespo, L.C., Henriques, S., Carvalho, R. & Gomes, P. (2012a) Determinants of spider species richness in coastal dunes along a gradient of

bioRxiv preprint doi: https://doi.org/10.1101/060400; this version posted June 23, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

mediterraneity. Insect Conservation and Diversity, 5: 127-137.

Carvalho, J.C., Cardoso, P. & Gomes, P. (2012b) Determining the relative roles of species replacement and species richness differences in generating beta-diversity patterns. Global Ecology and Biogeography, 21: 760-771.

Carvalho, J.C., Cardoso, P., Borges, P.A.V., Schmera, D. & Podani, J. (2013) Measuring fractions of beta diversity and their relationships to nestedness: a theoretical and empirical comparison of novel approaches. Oikos, 122: 825-834.

Coddington, J.A., Agnarsson, I., Miller, J.A., Kuntner, M. & Hormiga, G. (2009) Undersampling bias: the null hypothesis for singleton species in tropical arthropod surveys. Journal of Animal Ecology, 78: 573-584.

Collen, B., Loh, J., McRae, L., Whitmee, S., Amin, R. & Baillie, J.E.M. (2009) Monitoring change in vertebrate abundance: the Living Planet Index. Conservation Biology, 23: 317-327.

Crespo, L.C., Cardoso, P., Carvalho, R., Henriques, S.S., Rufino, A.C. (2009) Spiders (Arachnida: Araneae) from the Paúl de Arzila Natural Reserve (Portugal). Boletín de la Sociedad Entomologica Aragonesa, 44: 305-313.

Crespo, L.C., Silva, I., Borges, P.A.V. & Cardoso, P. (2013) Rapid biodiversity assessment, faunistics and description of a new spider species (Araneae) from Desertas Islands and Madeira (Portugal). Revista Ibérica de Aracnología, 23: 11-23.

Danielsen, F., Jensen, A.E., Alviola, P.A., Balete, D.S., Mendoza, M., Tagtag, A., Custodio, C. & Enghoff, M. (2005) Does monitoring matter? A quantitative assessment of management decisions from locally-based monitoring of protected areas. Biodiversity and Conservation, 14: 2633-2652.

Didham, R.K., Basset, Y. & Leather, S.R. (2010) Research needs in insect conservation and diversity. Insect Conservation and Diversity, 3: 1-4.

Didham, R.K., Leather, S.R. & Basset, Y. (2016) Circle the bandwagons – challenges mount against the theoretical foundations of applied functional trait and ecosystem service research. Insect Conservation and Diversity, 9: 1-3.

Faith, D.P. (1992) Conservation evaluation and phylogenetic diversity. Biological Conservation, 61: 1-10.

Faith, D.P. & Walker, P.A. (1996) Environmental diversity: on the best-possible use of surrogate data for assessing the relative biodiversity of sets of areas. Biodiversity and Conservation, 5: 399-415.

Graham, C. & Fine, P. (2008) Phylogenetic beta diversity: linking ecological and evolutionary processes across space and time. Ecology Letters, 11: 1265-1277.

Hilty, J. & Merenlender, A.M. (2000) Faunal indicator taxa selection for monitoring ecosystem health. Biological Conservation, 92: 185-197.

Hoffman, M., et al. (2010) The impact of conservation on the status of the world's vertebrates. Science, 330: 1503-1509.

Hoffman, M., Belant, J.L., Chanson, J.S., Cox, N.A., Lamoreux, J., Rodrigues, A.S.L., Schipper, J. & Stuart, S.N. (2011) The changing fates of the world's mammals. Philosophical Transactions of the Royal Society B, 366: 2598-2610.

Hortal, J. & Lobo, J.M. (2005) An ED-based protocol for optimal sampling of biodiversity. Biodiversity and Conservation, 14: 2913-2947.

Laliberté, E. & Legendre, P. (2010) A distance-based framework for measuring functional diversity from multiple traits. Ecology, 91: 299-305.

bioRxiv preprint doi: https://doi.org/10.1101/060400; this version posted June 23, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Leakey, R.E. & Lewin, R. (1996) The Sixth Extinction: Biodiversity and Its Survival. Science Masters.

Legendre, P. (2014) Interpreting the replacement and richness difference components of beta diversity. Global Ecology and Biogeography, 23: 1324-1334.

Loh, J., Green, R.E., Ricketts, T., Lamoreux, J.F., Jenkins, M. et al (2005) The Living Planet Index: using species population time series to track trends in biodiversity. Philosophical Transactions of the Royal Society of London B, 360: 289-295.

Longino, J.T. & Colwell, R.K. (1997) Biodiversity assessment using structured inventory: capturing the ant fauna of a tropical rain forest. Ecological Applications, 7: 1263-1277.

Muelelwa, M.I., Foord, S.H., Dippenaar-Schoeman, A.S. & Stam, E.M. (2010) Towards a standardized and optimized protocol for rapid assessments; spider species richness and assemblage composition in two savanna vegetation types. African Zoology, 45: 273-290.

Noss, R.F. (1990) Indicators for monitoring biodiversity: a hierarchical approach. Conservation Biology, 4: 355-364.

Pavoine, S., Vallet, J., Dufour, A. B., Gachet, S., & Daniel, H. (2009). On the challenge of treating various types of variables: application for improving the measurement of functional diversity. Oikos, 118: 391-402.

Pereira, H.M., Belnap, J., Brummitt, N., Collen, B., Ding, H., Gonzalez-Espinosa, M., Gregory, R.D., Honrado, J., Jongman, R.H.G., Julliard, R., McRae, L., Proença, V., Rodrigues, P., Opige, M., Rodriguez, J.P., Schmeller, D.S., Van Swaay, C. & Vieira, C. (2010). Global biodiversity monitoring. Frontiers in Ecology and the Environment, 8: 459-460.

Pereira, H.M., Ferrier, S., Walters, M., Geller, G.N., Jongman, R.H.G., Scholes, R.J., Bruford, M.W., Brummitt, N., Butchart, S.H.M., Cardoso, A.C., Coops, N.C., Dulloo, E., Faith, D.P., Freyhof, J., Gregory, R.D., Heip, C., Höft, R., Hurtt, G., Jetz, W., Karp, D.S., McGeoch, M.A., Obura, D., Onoda, Y., Pettorelli, N., Reyers, B., Sayre, R., Scharlemann, J.P.W., Stuart, S.N., Turak, E., Walpole, M. & Wegmann, M. (2013) Essential biodiversity variables. Science, 339: 277-278.

Petchey, O.L. & Gaston, K.J. (2002) Functional diversity (FD), species richness and community composition. Ecology Letters, 5: 402-411.

Petchey, O.L. & Gaston, K.J. (2006) Functional diversity: back to basics and looking forward. Ecology Letters, 9: 741-758.

R Core Team (2016) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Accessed 27.05.2016. URL http://www.R-project.org/.

Rao, C.R. (1982) Diversity and dissimilarity coefficients: a unified approach. Theoretical Population Biology, 21: 24-43.

Ricotta, C., Bacaro, G., Marignani, M., Godefroid, S. & Mazzoleni, S. (2012) Computing diversity from dated phylogenies and taxonomic hierarchies: does it make a difference to the conclusions? Oecologia, 170: 501-506.

Schilthuizen, M., Vairappan, C.S., Slade, E.M., Mann, D.J. & Miller, J.A. (2015) Specimens as primary data: museums and 'open science'. Trends in Ecology and Evolution, 30: 237-238.

Stork, N.E., Samways, M.J. & Eeley, H.A.C. (1996) Inventorying and monitoring biodiversity. Trends in Ecology and Evolution, 11: 39-40.

bioRxiv preprint doi: https://doi.org/10.1101/060400; this version posted June 23, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Ståhls, G., Vujic, A., Petanidou, T., Cardoso, P., Radenkovic, S., Acanski, J., Pérez-Bañon, C., Rojo, S. (2016) Phylogeographic patterns of *Merodon* hover flies in the Eastern Mediterranean region: revealing connections and barriers. Ecology and Evolution, 6: 2226-2245.

Tilman, D., Reich, P.B., Knops, J., Wedin, D., Mielke, T. & Lehman, C. (2001) Diversity and productivity in a long-term grassland experiment. Science, 294: 843-845.

Van Swaay, C., Regan, E., Ling, M., Bozhinovska, E., Fernandez, M., Marini-Filho, O.J., Huertas, B., Phon, C.-K., Korösi, A., Meerman, J., Peer, G., Uehara-Prado, M., Sáfián, S., Sam, L., Shuey, J., Taron, D., Terblanche, R., and Underhill, L. (2015) Guidelines for Standardised Global Butterfly Monitoring. Group on Earth Observations Biodiversity Observation Network, Leipzig, Germany. GEO BON Technical Series 1, 32pp.

Villéger, S., Miranda, J.R., Hernandez, D.F. & Mouillot, D. (2012) Low functional β-diversity despite high taxonomic β-diversity among tropical estuarine fish communities. PLoS ONE, 7: e40679.

Webb, C.O., Ackerly, D.D., McPeek, M.A. & Donoghue, M.J. (2002) Phylogenies and community ecology. Annual Review of Ecology, Evolution and Systematics, 33: 475-505.

World Spider Catalog (2016) World Spider Catalog. Natural History Museum Bern, online at http://wsc.nmbe.ch, version 17.0, accessed on 23.06.2016.

Yoccoz, N.G., Nichols, J.D. & Boulinier, T. (2001) Monitoring of biological diversity in space and time. Trends in Ecology & Evolution, 16: 446-453.