For all audiences: Incorporating immature stages into standardised inventories of mega-diverse groups has a major impact on our understanding of biodiversity patterns

Marc Domènech[1], Owen S. Wangensteen[2], Alba Enguídanos[1], Jagoba Malumbres-Olarte[3], and Miquel Arnedo[1]

[1]University of Barcelona
[2]University of Tromsø
[3]University of the Azores

August 19, 2021

Abstract

Because of their challenging taxonomy, arthropods are traditionally underrepresented in biological inventories and monitoring programs. However, arthropods are the largest component of biodiversity, and no assessment can be considered informative without including them. Arthropod immature stages are often discarded during sorting, despite frequently representing more than half of the collected individuals. To date, little effort has been devoted to characterising the impact of discarding non-adult specimens on our diversity estimates. Here, we use a metabarcoding approach to analyse spiders from white oak communities in the Iberian Peninsula collected with standardised protocols, to assess (1) the contribution of juvenile stages to local diversity estimates, and (2) their effect on the diversity patterns inferred across communities. We further investigate the ability of metabarcoding to inform on abundance. We obtained 363 and 331 species as adults and juveniles, respectively. Species represented only by juveniles represented an increase of 35% with respect to those identified from adults in the whole sampling. Differences in composition between communities were greatly reduced when immature stages were taken considered, especially across latitudes. Moreover, our results revealed that metabarcoding data are to a certain extent quantitative, but some sort of taxonomic conversion factor may be necessary to provide accurate informative estimates. Although our findings do not question the relevance of the information provided by adult-based inventories, they also reveal that juveniles provide a novel and relevant layer of knowledge that, especially in areas with marked seasonality, may influence our interpretations, providing more accurate information from standardised biological inventories.

For all audiences: Incorporating immature stages into standardised inventories of mega-diverse groups has a major impact on our understanding of biodiversity patterns

Marc Domènech[1], Owen S. Wangensteen[2], Alba Enguídanos[1], Jagoba Malumbres-Olarte[3,4] & Miquel A. Arnedo[1]

[1] Department of Evolutionary Biology, Ecology and Environmental Sciences & Biodiversity Research Institute (IRBio), Universitat de Barcelona, Av. Diagonal 643, E-08028, Barcelona, Spain.

[2] Norwegian College of Fishery Science, University of Tromsø, The Arctic University of Norway, Norway.

[3] cE3c – Centre for Ecology, Evolution and Environmental Changes / Azorean Biodiversity Group and Universidade dos Açores – Faculty of Agrarian and Environmental Sciences, Rua Capitao Joao d'Avila, Sao Pedro, 9700-042 Angra do Heroismo, Terceira, Açores, Portugal.

[4] LIBRe – Laboratory for Integrative Biodiversity Research, Finnish Museum of Natural History, University of Helsinki, P.O.Box 17 (Pohjoinen Rautatiekatu 13), 00014 Helsinki, Finland

Correspondence : Marc Domènech, Department of Evolutionary Biology, Ecology and Environmental Sciences & IRBio, Universitat de Barcelona, Barcelona, Spain. Email: mdomenan@gmail.com

ABSTRACT

Because of their challenging taxonomy, arthropods are traditionally underrepresented in biological inventories and monitoring programs. However, arthropods are the largest component of biodiversity, and no assessment can be considered informative without including them. Arthropod immature stages are often discarded during sorting, despite frequently representing more than half of the collected individuals. To date, little effort has been devoted to characterising the impact of discarding non-adult specimens on our diversity estimates.

Here, we use a metabarcoding approach to analyse spiders from white oak communities in the Iberian Peninsula collected with standardised protocols, to assess (1) the contribution of juvenile stages to local diversity estimates, and (2) their effect on the diversity patterns inferred across communities. We further investigate the ability of metabarcoding to inform on abundance. We obtained 363 and 331 species as adults and juveniles, respectively. Species represented only by juveniles represented an increase of 35% with respect to those identified from adults in the whole sampling. Differences in composition between communities were greatly reduced when immature stages were taken into account, especially across latitudes. Moreover, our results revealed that metabarcoding data are to a certain extent quantitative, but some sort of taxonomic conversion factor may be necessary to provide accurate informative estimates.

Although our findings do not question the relevance of the information provided by adult-based inventories, they also reveal that juveniles provide a novel and relevant layer of knowledge that, especially in areas with marked seasonality, may influence our interpretations, providing more accurate information from standardised biological inventories.

Keywords : Araneae, Diversity, DNA barcoding, Iberian Peninsula, Metabarcoding, Spiders

Running title: Incorporating juveniles in diversity estimates

INTRODUCTION

Global human activity is altering the species richness and abundance of biological communities (Stuart Chapin III et al 2000; Socolar, Valderrama-Sandoval, & Wilcove, 2019; Barlow et al., 2016). These disturbances are accelerating the invasion rate of exotic species (Hulme, 2009) and driving numerous species to extinction, sometimes even before they are described, in what has been referred to as the*sixth mass extinction* (Dirzo et al, 2014; Barnosky et al, 2011). Bioinventories and early detection methods for monitoring ecosystem changes are essential to identify and tackle unanticipated threats to biodiversity (Telfer et al 2015; Barnosky et al 2013). However, sampling and identifying highly abundant and diverse groups such as arthropods is a daunting task. Several methods have been devised to overcome this limitation. For instance, rapid biodiversity assessment protocols are fast yet efficient sampling strategies specifically designed to retrieve the greatest amount of information from a particular area, minimising the number and length of sampling periods (Oliver and Beattie 1996).

Arthropods represent the largest and most abundant component of animal biodiversity. Therefore, no biodiversity monitoring program can be considered credible unless it takes them into account (Taylor and Doran 2001). Moreover, because of their high reproductive rates and short generation times, arthropods have the potential to inform on biodiversity changes at finer spatial and shorter temporal scales than vertebrates (Kremen et al 1993, Yen & Butcher 1997). However, the poor taxonomic knowledge of many arthropods limits their use as bioindicators, and they are frequently underrepresented in biodiversity assessments and conservation programs (Cardoso, Borges, Triantis, Ferrández, & Martín, 2012). On the other hand, recent studies suggest that arthropod populations, which play a fundamental role in ecosystem functioning, are declining at an alarming rate (Leather 2018).

Because of their rapid and divergent evolution, male copulatory apparatus and, to a lesser degree, female external reproductive systems are the main structures for species identification across most arthropod groups (Eberhard, 1985). Spiders are no exception. Morphological taxonomic identification is almost exclusively based on genitalic characters, i. e., the structure of the copulatory bulb in males and the vulva and epigyne (external modifications of the genital area) in females. These features are only visible in the last moult, which makes immatures difficult or impossible to identify at species level (Dobyns, 1997; Coddington, Young, & Coyle, 1996). Thus, in most inventories and diversity studies immature stages are discarded during sorting. However, immature specimens may account for between 40% and 70% of the collected specimens in biodiversity surveys (Soukainen et al, 2020, Malumbres-Olarte et al, 2020a, Malumbres-Olarte, Cardoso, & Crespo, 2019; Cardoso, Silva, Oliveira, & Serrano, 2004; Russell-Smith & Stork 1995, Silva 1996), or even up to 94% in extreme cases (Kuntner & Baxter, 1997). Disregarding immatures may significantly influence the inference of the temporal and spatial patterns of biodiversity, so their incorporation is desirable to obtain reliable estimates of diversity in short-term sampling protocols (Toti, Coyle, & Miller, 2000; Sorensen, Coddington, & Scharff, 2002).

It is known that the life cycles of different groups of spiders differ in the number of generations per year and in the time of the year when they are present as adults or juveniles (Aitchison, 1984; Nadal, Achitte-Schmutzler, Zanone, Gonzalez, & Avalos, 2018), even within the same species. For example, some wolf spiders are known to have annual life cycles maturing as adults in March-April and reproducing in May-June, although they can also have two clutches in the same year depending on the weather conditions (Rádai, Kiss, & Samu, 2017). Since rapid biodiversity assessment protocols are usually conducted once, they provide a "photograph" of the species present in a certain area at one particular time, so considering only adult individuals would completely dismiss all the species that are present as immature stages in that particular time of the year. To our knowledge, only one study (Norris, 1999) has partially addressed the effect in diversity estimates of incorporating juvenile spider stages. This study, which included only a few species that could reliably be morphologically identified at immature stages, already pointed out that numerous species were only found as juveniles and that relative abundances changed drastically when these immature stages were taken into account.

The use of DNA-based approaches for species identification, e.g., DNA barcoding (Hebert, Ratnasingham, & DeWaard, 2003), ease the identification of immature stages. DNA barcode sequences of immature individuals can be assigned to species through comparison to reference databases containing barcodes of adult-based morphologically identified species (Richard et al, 2010; Meiklejohn, Wallman, & Dowton, 2012). However, this technique still poses some drawbacks. For example, it requires manually extracting and amplifying each specimen individually, which for large samplings with hundreds or thousands of juveniles can be very time-consuming. Moreover, the economic cost of extracting, amplifying and sequencing such a large amount of samples would also be considerable.

DNA metabarcoding is a more recently developed molecular technique consisting in the automated identification of multiple species from a single bulk sample containing entire or partial organisms or from environmental samples (water, soil, etc.) containing remains of DNA (e.g., Bohmann et al., 2014; Yu et al., 2012; Morinière et al., 2016). This approach represents a clear advantage with respect to DNA barcoding, as it allows the simultaneous processing of a large number of specimens at once, greatly reducing the workload and processing time. In addition, it is more cost efficient for large numbers of specimens, as the number of sequences obtained from a single metabarcoding run is in the order of millions (Sales et al., 2020; Watts et al., 2019). The downside of using this approach is that individual specimens cannot be traced back or are sometimes even lost in the process of preparing the bulk sample, making it impossible to revise the voucher specimens if interesting sequences were found.

Another potential drawback of the use of metabarcoding for biodiversity assessment is its presumed inability to provide abundance information. To what extent the number of sequence reads of a certain taxon correctly represents its abundance or biomass in the sample has been a matter of much debate. Several studies have specifically addressed this issue (Elbrecht and Leese, 2015; Piñol, Mir, Gomez-Polo, & Agustí, 2015; Lamb

et al., 2019; Deagle et al., 2019), but the answer remains inconclusive. While some studies consider the quantitative power of metabarcoding limited (Elbrecht and Leese, 2015; Piñol et al., 2015), others found metabarcoding to give an accurate estimate of a taxon's abundance under certain conditions (Ratcliffe et al., 2020) or by applying correction factors that may vary among taxa (Kennedy et al., 2020; Thomas, Deagle, Eveson, Harsch, & Trites, 2016), or have even used it to quantitatively analyse dietary data (Soinien et al., 2015). One explanation for these different conclusions may be that the range of concentrations analysed varies considerably across studies, as suggested by Deagle et al. (2019). While a positive relationship between number of reads and biomass is commonly found, only a certain part of the variation in the number of reads seems to be explained by differences in the biomass in the sample, while the rest of the variation seems to be due to factors such as primer specificity (Elbrecht and Leese, 2015), different extraction success between different tissues or species (Schiebelhut, Abboud, Gómez-Daglio, Swift, & Dawson, 2016) or the efficiency of the blocking primers of predator DNA in the case of diet studies (Piñol et al., 2015).

Aim

Here we aim to evaluate the impact of including juveniles of spiders into community diversity estimates. Specifically, we first quantify the additional diversity that immature spiders contribute to diversity estimates. Secondly, we investigate how the inclusion of immature specimens affects biodiversity patterns across communities and identify the possible ecological factors responsible for such patterns. Finally, we assess the ability of metabarcoding to recover abundance information from bulk samples. Our results provide important insights into the relevance of considering all different life stages in rapid biodiversity assessment protocols, which are essential to efficiently monitor ecosystem changes, and will contribute to refine the use of metabarcoding approaches as efficient alternatives to traditional, morphology-based standardised biological inventorying and monitoring schemes.

MATERIAL AND METHODS

Sample collection and sorting

We collected the specimens using the standardised sampling protocol COBRA (Cardoso, 2009) in May-June 2013 and 2014 (Crespo et al., 2018; Malumbres-Olarte, Crespo, Domènech, Moya-Laraño, & Arnedo, 2020b). The sampling design included 16x1ha plots distributed in white oak forests across six National Parks of the Iberian Peninsula, namely Aigüestortes i Estany de Sant Maurici (PA), Ordesa y Monte Perdido (PO), Picos de Europa (PP), Monfragüe (PM), Cabañeros (PC) and Sierra Nevada (PS) (Fig. 1). We used semi-quantitative methods that combined 12 man-hours of timed direct capture, beating and sweeping with 48 pitfall traps active for two weeks in every plot.

We sorted adult and immature specimens and we identified them under a ZEISS Stemi 2000 stereomicroscope. We identified adults to species level and juveniles to family level. All the immature individuals of each plot were weighted separately for each family using an analytical balance. We placed the specimens in absorbent paper for 30 minutes before weighing them to allow the remaining alcohol in the bodies to evaporate.

DNA extraction, amplification and sequencing

Representative DNA barcode sequences for the cytochrome c oxidase subunit I (COI) (˜658 bp) of species captured as adults were available from a previous study (Crespo et al. 2018) (see Table S1 of Supporting information and Data Accessibility).

For each plot, we homogenised all the collected juveniles with the help of liquid nitrogen. We obtained two extraction replicates from each homogenised plot sample, and we extracted a fraction of 0.3 g from each replicate using a PowerSoil DNA Isolation Kit (QIAGEN, Valencia, CA, USA). We added one negative (distilled water) and one positive control, a specimen of the cobweb spider *Simitidion simile* (C. L. Koch, 1836), in the extraction protocol. These controls were included in the batch, processed and sequenced along with the rest of the samples. We cleaned and sterilised all the equipment with diluted sodium hypochlorite between successive sample extractions. We amplified the COI "Leray fragment" of 313 bp using the

degenerate primer set Leray-XT (Wangensteen, Palacín, Guardiola, Turon, 2018). This set includesthe reverse primer jgHCO2198 5'-TAIACYTCIGGRTGICCRAARAAYCA-3' (Geller, Meyer, Parker, & Hawk, 2013) and the forward primer mlCOIintF-XT 5'-GGWACWRGWTGRACWITITAYCCYCC-3', modified from the mlCOIintF primer (Leray et al., 2013). Each primer pair included twin 8-bp sample tags (the same tag in the forward and reverse primers) and a lead of 2–4 random Ns in the 5' end for increasing sequence variability of the library. The PCR mix included 10 µl AmpliTaq Gold 360 Master mix (Applied Biosystems, Foster City, CA, USA), with 1 µl of each 5 µM forward and reverse primers, 0,16 µl of bovine serum albumin, 2 µl of DNA template and DNase-Free water to adjust the volume up to 20 uL per sample. The PCR profile included 10 min at 95 °C, 35 cycles of 94 °C 1 min, 45 °C 1 min and 72 °C 1 min, and 5 min at 72 °C. We used two PCR replicates of each extraction replicate in the study, giving a total of 4 replicates per plot (except for plots PA1 and PS1, for which we obtained 3 replicates as they were the first to be processed and served as a test). We evaluated the quality of amplifications by electrophoresis in 1% agarose in Tris-borate-EDTA buffer and stained with GelRed® Nucleic Acid Gel Stain (Biotium, Hayward, California, USA). We pooled all PCR products by equal volume (including two PCR-negative controls and one PCR-positive control) and purified them using a MinElute PCR Purification Kit (Qiagen, Valencia, CA, USA). Three µg of the purified pool (determined by Qubit fluorometric quantitation dsDNA BR Assay Kit, Thermo Fisher Scientific, Waltham, Massachusetts, USA) were used to build a library using the NextFlex PCR-free DNA-seq kit (Perkin-Elmer, Waltham, Massachusetts, USA). The multiplexed library was sequenced in an Illumina MiSeq with a V3 2x250 bp paired-end partial run at the University of Salford, UK.

Bioinformatics

We conducted the bioinformatic analyses using the Obitools metabarcoding package (Boyer et al., 2016). We aligned the paired-end reads using the command *illuminapairedend* . We selected sequences with alignment quality scores bigger than 40 and we demultiplexed the aligned dataset and removed the primer sequences with *ngsfilter* . We also filtered out sequences containing ambiguous bases. We then used *Obiuniq* to dereplicate the reads (grouping all identical sequences) while keeping track of their abundances, and we also removed chimeric sequences using the *uchime_denovo* algorithm in VSearch (Rognes, Flouri, Nichols, Quince & Mahé, 2016). We used the step-by-step aggregation clustering algorithm implemented in Swarm 2.1.13 (Mahé, Rognes, Quince, de Vargas, & Dunthorn, 2015) to cluster the sequences into Molecular Operational Taxonomic Units (MOTUs). For making adult (morphology and DNA barcode data) and juvenile (metabarcoding data) clustering comparable, we combined the sequences from both life stages before running the Swarm clustering algorithm. In the case of adults, we kept only the segment of the original COI sequences matching the Leray COI fragment. To prevent the program from discarding adult sequences as singletons, we artificially increased their initial abundance to 50,000 reads. We set a distance value of d = 13 for the clustering algorithm, which has been shown to be the optimal value for discriminating intra and interspecific divergences, that is, to approximate MOTUs to species-level clusters, in a wide range of eukaryotic systems (Wangensteen & Turon, 2017; Kemp et al., 2019; Siegenthaler et al, 2019; Garcés-Pastor et al., 2019; Antich, Palacín, Wangensteen & Turón, 2021). The species present as adults whose sequences were clustered together by Swarm (nine pairs, one triad and one tetrad) were also treated as single entities in downstream analyses with juveniles. After removing the singletons, we performed the taxonomic assignment of the representative sequences of each MOTU (seeds) using Ecotag (Boyer et al., 2016). We built the local reference sequence database required by Ecotag, combining our sequences of adult spiders with sequences retrieved from the BOLD database (Ratnasingham & Hebert, 2007) and the EMBL repository (Kulikova et al., 2004). Ecotag (Boyer et al. 2016) uses a phylogenetic assignment protocol, based on the NCBI taxonomy tree, to assign sequences to the last common ancestor of the most closely related sequences in the local reference database. This approach does not require establishing arbitrary identity thresholds for every taxonomic rank (Bakker et al., 2019).

We filtered out putative contaminants of the resulting database by retaining only the MOTUs assigned to the order Araneae. After the taxonomic assignment made by Ecotag, we manually checked if there were better, more recent matches in BOLD or NCBI, and we updated the identification of those MOTUs for which better matches were found. We discarded as contaminants 16 MOTUs with low numbers of reads that corresponded

to a checklist of non-iberian species that had been analysed in other studies conducted in the same lab. We used the LULU algorithm (Frøslev et al., 2017) to remove the MOTUs corresponding to pseudogenes. We also built a COI tree using the seed sequence of every MOTU and the COI sequence of the adult specimens to help allocate unassigned MOTUs to specific families, genera or species. We inferred the tree by Maximum Likelihood using IQ-TREE v.1.6 (Nguyen, Schmidt, von Haeseler, & Minh, 2015). We partitioned positions by codon and assigned an unlinked GTR model to each partition, and we assessed branch support by means of 1,000 ultrafast bootstrap approximation replicates (Minh, Nguyen, & von Haeseler, 2013; Hoang, Chernomor, von Haeseler, Minh & Vinh, 2018). Analyses were run remotely at the CIPRES Science Gateway (Miller, Pfeiffer, & Schwartz, 2010). All the replicates of each plot were added up. All the MOTU's with less than five total reads were discarded. Also, for a MOTU to be counted as present in a plot, we required at least five reads in the plot and detection of the MOTU in at least two of the replicates of the plot.

Delimitation of adult and juvenile clusters

For the MOTUs that could not be identified to species level, we used the best match to which the taxonomic assignment algorithm assigned that MOTU, that is, the identifier of the specific sequence in the database which was the most similar to the seed sequence of the MOTU. If the best match of two unidentified MOTUs was the same, they were collapsed and were treated as the same taxon. For comparative purposes, we also analysed the results in two additional alternative ways: using a "splitter" approach (every unidentified MOTU as a different species) and using a "lumper" approach (all the unidentified MOTUs of the same genus/family considered as the same species). However, using either of those approaches only translated into minor differences in the results with respect to the "best match" approach.

Evaluation of the effect of including juveniles on community patterns

We checked the completeness of the sequenced replicates by means of rarefaction curves, plotting the number of MOTUs per replicate against an increasing number of reads (Fig. S1, Supporting information). To find out if there were significant differences in the similarity patterns among communities when including juveniles, we performed a non-metric multidimensional scaling (NMDS) analysis based on the community composition. We performed these analyses with presence/absence data both with the information on adults alone and with adults and juveniles together. For this, we used the *metaMDS* function in the package "vegan" (Oksanen et al, 2019) in R (R Core Team, 2020). We applied a Mantel test to assess the correlation and significance between the distance matrices obtained from both approaches. We also applied an ANalysis Of SIMilarity (ANOSIM) analysis to test if differences in species composition between northern and southern parks were equally recovered by adults and adults+juveniles approaches.

Assessment of species abundance from metabarcoding data

To assess the level of variation in the number of reads explained by the proportion of a certain taxa in the sample, we calculated the proportion of reads or relative read abundance (RRA), the proportion of biomass and the proportion of individuals of every family in every plot. We used the*betareg* function in the R package "betareg v1.1" (Cribari-Neto & Zeileis, 2010) to apply beta regression models to each family present in at least 10 plots. Two models were applied to each family, one for the RRA as a function of the percentage of mass the family represents in the sample, and one for the RRA as a function of the percentage of abundance. We applied the models to each family separately due to the non-independence of percentages in a sample, and we used 10 as the minimum presence in plots following the one-in-ten rule tested in other models (Peduzzi, Concato, Feinstein, & Holford, 1995). Beta regression models are designed for response variables with proportional data between 0 and 1. We calculated the adjusted R^2 values and determined the global goodness of fit for each model. Only significant (p-value < 0.05) models with a pseudo-$R^2 > 0.5$ and randomly distributed residuals were considered (Yellareddygari, Pasche, Taylor, Hua, & Gudmestad, 2015).

RESULTS

Sequencing results

We retrieved DNA barcodes of 368 out of the 376 species represented as adults in the sampling plots. After the

cleaning, filtering, and chimaera removal process, the sequencing of juvenile pool samples generated a total of 15,805,993 sequence reads and 3,839,513 unique sequences. After manually adding the sequences of the adults, the Swarm algorithm grouped all sequences in 4,668 non-singleton MOTUs, of which 140 contained exclusively adult artificial sequences. After the curation process (retaining only Araneae, removing non-iberian taxa, and removing artificially added adult sequences), the final dataset for juveniles consisted of 9,956,432 reads distributed in 1,343 MOTUs. The process of removing pseudogenes eliminated 455 MOTUs and left 888 final juvenile MOTUs (final dataset in Table S2, Supporting information). The sequencing of three of the four replicates of plot PM1 was unsuccessful so we omitted these samples from subsequent analyses.

The steps involved in refining the taxonomic assignment of these MOTUs (merging the species that were found to be indistinguishable by the LERAY COI fragment, checking online genetic databases and the COI tree to improve taxonomic assignments, and assigning the remaining unassigned MOTUs to their best match in the reference library) assigned the 888 MOTUs to 524 different species. Filtering out MOTUs with less than 5 reads left 411 MOTUS that, when collapsed by their taxonomic assignment, corresponded to 350 different species. Finally, considering only MOTUs with at least 5 replicates in a plot and presence in at least 2 of its replicates as present in that plot left 331 different species.

Rarefaction curves (Fig. S1, Supporting information) showed that almost every plot reached a plateau in all or some of its replicates. The exceptions were the only remaining replicate of PM1 and two of the replicates of PS2, all of which yielded very low numbers of reads.

Delimitation of adult and juvenile clusters

After collapsing the species that were indistingible by the COI fragment, the final values of species richness for adults and juveniles were 363 and 331, respectively. The combination of both lifestages yielded a total of 491 different species. Of those, 160 were found exclusively as adults, 128 exclusively as juveniles and 203 as both life stages. The addition of the species found only as juveniles represented a 35% increase with respect to all the species found as adults. Excluding PM1 (whose sequencing was unsuccessful) the number of matching species varied from 24 (in PM2) to 44 (in PP4) (Fig. 2). As for the percentages, PM2 was the plot where juveniles provided the greatest addition of richness (169% more species than only with adults), followed by PC2 (162%), PC3 (119%), PC4 (117%), PP3 (112%), PC1 (102%), PP1 (92%), PS2 (83%), PS1 (73%), PO1 (68%), PA1 (58%), PP2 (56%), PO2 (54%), PP4 (47%) and PA2 (44%).

The number of species recovered as juveniles was higher than that of adults in every park (5.4% higher in PO, 8.2% higher in PP, 17.6% higher in PA, 29.8% higher in PS, 45.9% higher in PC and 69.1% higher in PM). Most of the families showed a similar or identical species richness in adults and juveniles, but some specific families showed clear differences (Fig. 3). For example, the richness of juvenile orb-weavers (Araneidae) was much greater than the richness of adults in all the parks, and a similar trend (although to a lesser extent) was also observed in cob-weaving spiders (Theridiidae) and the sit-and-wait hunting families Philodromidae and Thomisidae crab spiders. Interestingly, for sheet-weaving spiders (Linyphiidae) the number of species represented as adults was greater than that of juveniles in the northern parks (as much as twice as high in PP) but the opposite trend was observed in the southern parks. While the number of sheet-weaving spider MOTUs collected as juveniles remained very similar across all parks, ranging from 12 to 17, the number of species recovered as adults decreased abruptly from the northern parks (16, 24 and 34) to the southern parks (7, 10 and 11).

Evaluation of the effect of including juveniles on community patterns

In the adult-based NMDS (Fig. 4a) the distances between the communities of the same park were generally lower than the distances between the parks. This ordination showed a clear separation between northern communities and southern communities along the first component. The NMDS performed including individuals of all life stages (Fig. 4b) did not show a clear north-south distinction, and the parks were less homogeneous than in the NDMS based on adults. The Mantel test revealed that the correlation between the two distance matrices was significantly low ($r = 0.336$, $p = 0.004$). There were significant differences between the species

composition of northern and southern communities with the dataset containing only adult species (R=0.944, p=0.0005), but not with the dataset including juveniles (R=0.136, p=0.078).

Assessment of species abundance from metabarcoding data

After selecting the spider families that were present in at least 10 communities, we built beta regression models with data for the remaining 14 families. Seven of the families provided appropriate models for the proportion of weight, and six families did so for the models relating proportion of reads and proportion of individuals (Fig. 5). In the families for which an adequate model could not be fitted, there was still a positive correlation between the two variables.

Although the relation between RRA and weight or abundance was positive in all models, its slope varied across spider families. Even in the three families where both models had a $p > 0.05$, the two curves were almost overlapping in two of the families (Lycosidae and Salticidae) but rather different in the third (Clubionidae). Also, in some families the observations were consistently above or below the 1:1 line. In the plots linking RRA and abundance of individuals, funnel-web (Agelenidae) and cob-weaving spiders (Theridiidae) were above the 1:1 line, while ghost spiders (Anyphaenidae) and ground-dwellers (Clubionidae) were mostly below. In the plots linking RRA and mass, Clubionidae ground-dwellers, Philodromidae crab spiders and Theridiidae cob-weavers were above the 1:1 line, while Agelenidae funnel-weavers, Araneidae orb-weavers, Gnaphosidae ground-dwellers and Lycosidae wolf spiders were mostly below.

DISCUSSION

An important component of biodiversity at a given time is mostly represented by juveniles

Immature specimens represented 59% of all the captured specimens on average, ranging between 39% - 76% per plot. This result is similar to those obtained in other studies (Soukainen et al, 2020, Malumbres-Olarte et al, 2020a, Malumbres-Olarte et al, 2019; Cardoso et al 2004; Russell-Smith & Stork 1995, Silva 1996), which ranged from 40% to 70%, and offers a first insight on the relevance of juvenile stages in spider inventories and the conclusions derived from them.

Overall, the number of species estimated from juveniles by metabarcoding was lower than that of adults. This could suggest that, at the time of the samplings, there were fewer species in the juvenile stage than in the adult stage. However, this finding could well be an artifact related to the sampling methods, as immature spiders are smaller than adults and may be more difficult to detect by direct sampling techniques. The combination of both life stages yielded a total of 491 different species, 35% higher than the richness we obtained considering only adults in the whole sampling. The degree to which immature stages contributed to diversity, however, was not constant across all our plots. In almost half of the individual samplings, the total number of species including juveniles more than doubled the richness obtained with adults. These results indicate that a large part of the diversity may be ignored by spider bioinventories that are performed exclusively on adults. Interestingly, the contribution of juveniles was higher in southern parks than in northern ones. Although knowing the reasons behind this pattern would require further study, we suspect the difference may be related to a phenological delay between both latitudes.

Although most families recovered a similar species richness in adult and in juvenile stages, some had important differences (Fig. 3). In the case of araneids, the fact that the number of captured juveniles was almost twice as high as the number of adults may have made the number of species captured as immatures greater because it increased the chances of sampling additional species for this stage. The additional diversity found only as juvenile spiders may also be related to the phenology of this family. Larger orb-weavers (araneids) mature in autumn in temperate zones, while smaller species tend to mature earlier (Levi, 1973). This observation fully matches our findings, as most of the araneids that were found exclusively as juveniles are large species of orb-weavers, such as *Argiope lobata* Pallas, 1772, *Larinioides patagiatus* (Clerck, 1757) or several *Araneus* species.

Also interesting is the case of linyphiids sheet-weavers, whose adult species richness was higher than juvenile richness in the northern parks (as much as twice as high in PP), but lower in the southern parks. Linyphiids

are known to be much more diverse in temperate regions than in the subtropics and tropics (Cardoso, Pekár, Jocqué & Coddington, 2011). We indeed found more total species in northern parks, which have a more continental climate, than in southern parks, with a significantly drier and warmer climate. The differences in adult and juvenile richness between north and south, however, may indicate the existence of different predominant phenologies within linyphiids at both latitudes.

Although our data had a large taxonomic scope (order level), there are still some caveats to our analyses. All the MOTUs that could not be assigned to any nominal species could provide important additional information. Nevertheless, it is expected that completeness of reference databases will increase in coming years and the accuracy of taxonomic assignments in metabarcoding studies will improve.

Identifying juveniles changes our interpretation of similarities among communities

The Mantel test revealed that there are substantial differences in the species compositions of communities when considering only adult specimens or individuals of all life stages, and the ANOSIM showed that taking immature specimens into account reduced the differences in species composition between communities at different latitudes. The NMDS plots also revealed differences in community similarities between the two approaches. While in the NMDS with only adults most communities were more similar to other communities within the same park than to communities from different parks, this pattern was much less pronounced in the NMDS that includes juveniles. In addition, the marked distinction between communities at higher latitudes and communities at lower latitudes found with adults only was not recovered when considering all the individuals. This lower structure in the NMDS based on specimens of all life stages suggests that phenological differences among communities may lead to see community compositions as more different than what they really are. In the time of the sampling, a species might be present as adult in a certain community and still as juvenile in another community, so by taking into account exclusively adult stages we might be inadvertently accentuating the differences among community compositions.

Given that seasonality and marked phenologies with different life stages present at different times of the year are very common across many animal groups (Scott & Epstein, 1987; Jakob, Poizat, Veith, Seitz, & Crivelli, 2003; Lazaridou-Dimitriadou & Sgardelis, 1997), we suspect that the phenomenon that we describe here might also apply to inventories performed with other organisms. Studies on other animal taxa comparing alpha diversity estimates of a community using only individuals of a certain life stage or all individuals might help reveal if this trend is constant across different taxa.

Metabarcoding data may provide abundance information

As expected, and in accordance with other studies investigating the quantitative power of metabarcoding in spiders (Kennedy et al., 2020) or other taxa (Thomas et al., 2016; Deagle et al., 2018; Lamb et al., 2019; Krehenwinkel et al., 2017; Schenck, Geisen, Kleinbölting, & Traunspurger, 2019), our results indicate that the relative read abundance (RRA) of a taxon is positively related to its proportion in both weight and abundance. However, the strength of this relation is not constant across spider families, which dissipates the possibility of using a unique correction factor to derive abundance information from read counts for all taxa obtained in metabarcoding analyses. Indeed, similar studies have also found these differences in the factor linking RRA and individual abundance across spider families (Kennedy et al., 2020) or RRA and mass across different taxa in other animal groups (Thomas et al., 2016).

Interestingly, the observations of some of the families in the plots relating RRA to proportion in mass were consistently above or below the 1:1 line. Upon detailed inspection, the spider families mostly above the identity line corresponded to families with small juvenile individuals (between 0.9 and 2.4 mg in mass), while families mostly below the identity line corresponded to those with large juvenile individuals (between 5.5 and 18.7 mg). Families with an intermediate juvenile mass (2.5 to 3 mg) were not clearly above or below the 1:1 line. This suggests that taxa with small or large juvenile sizes might be respectively over or underrepresented by their reads counts with respect to their real weight proportion in the sample in metabarcoding analyses. Additional studies with spiders and other taxa would help determine if this is a consistent trend and, if so, if it is applicable to other groups apart from spiders.

We agree with previous studies stating that, albeit with caution and with a certain degree of uncertainty, using the RRA with the corresponding correction factors as a surrogate of the occurrence of a taxon in the sample still provides more precise information of the community composition than using presence/absence data (Lamb et al., 2019; Kennedy et al., 2020). However, these correction factors need to be developed individually for different taxa, for example using mock communities included as quantitative controls during metabarcoding (Lamb et al., 2019).

CONCLUSIONS

Our study suggests that incorporating immature stages of spiders in bioinventorying initiatives has a relevant effect on diversity estimates, because a considerable proportion of the species present as juveniles is not found among adults. This impact goes beyond simply modifying species richnesses, as it also alters the level of similarity among communities and how they compare to each other. These findings do not question the information provided by adult-based inventories, but add a novel yet relevant layer of knowledge previously overlooked that may influence some of the interpretations derived from biological inventories. Adding juvenile information to rapid biodiversity assessment protocols provides more accurate data regarding comparisons of community composition. Metabarcoding analysis of all stages present in a sample enables more effective monitoring strategies, and ultimately better-informed conservation decisions.

The proportion of reads obtained from metabarcoding for certain spider families was positively related to their proportion in weight and abundance in the sample, suggesting that metabarcoding data are to a certain extent quantitative. The strength of this relation, however, was not constant across families, as already reported in former studies. Nonetheless, the use of read counts appropriately transformed with taxon-specific correction factors as a proxy of the occurrence of a taxon in the sample could still provide more accurate information about the community composition than simple presence/absence data.

ACKNOWLEDGEMENTS

Collections were conducted under the corresponding permits kindly provided by the following individuals and institutions: Miguel M. de la Hoz (Picos de Europa), Elena Villagrasa (Ordesa), Maria Merced Aniz Montes (Aiguestortes), Angel Rodriguez Martin (Monfrague), Angel Gomez Manzaneque (Cabaneros), and Blanca Ramos Losada (Sierra Nevada). We are grateful to all the people that contributed to the samplings as well as to all the park rangers that supported us in the field.

This work was funded by the Organismo Autonomo de Parques Nacionales (OAPN #485/2012) (to MA). MD was supported by APIF PhD fellowship from the University of Barcelona. Additional support was provided by 2017SGR73 (to MA) from the Catalan Government.

AUTHOR CONTRIBUTIONS

MA, JMO and MD designed the study. AE and MD performed the laboratory work. OW, MD and JMO performed the data analyses. MD, JMO and MA drafted the manuscript with contributions of all authors. All authors have revised and approved the final manuscript.

DATA ACCESSIBILITY

- Fasta file with sequences of adult specimens: Dryad, https://doi.org/10.5061/dryad.g79cnp5q0.

- Fasta file with sequences of immature specimens: Dryad, https://doi.org/10.5061/dryad.z08kprrcv.

REFERENCES

Aitchison, C. W. (1984). The phenology of winter active spiders.*Journal of Arachnology* , *12* , 249–271.

Antich, A., Palacín, C., Wangensteen, O. S., & Turon, X. (2021) To denoise or to cluster? That is not the question. Optimizing pipelines for COI metabarcoding and metaphylogeography. *BMC Bioinformatics* ,*2021* , 22-177.

Bakker, J., Wangensteen, O. S., Baillie, C., Buddo, D., Chapman, D. D., Gallagher, A. J., Guttridge, T. L., Hertler, H.,& Mariani, S. (2019). Biodiversity assessment of tropical shelf eukaryotic communities via pelagic eDNA metabarcoding. *Ecology and evolution* , *9* , 14341–14355.

Barlow, J., Lennox, G. D., Ferreira, J., Berenguer, E., Lees, A. C., Mac Nally, R., Thomson, J. R., Ferraz, S. F. D. B., Louzada, J., Oliveira, V. H. F., Parry. L., Ribeiro De Castro Solar, R., Vieira, I. C. G., Aragaõ, L. E. O. C., Begotti, R. A., Braga, R. F., Cardoso, T. M., de Oliveira Jr, R. C., Souza, C. M., Moura, N. G., Nunes, S. S., Siqueira, J. V., Pardini, R., Silveira, J. M., Vaz-De-Mello, F. Z., Veiga, R. C. S., Venturieri, A., & Gardner, T. A. (2016). Anthropogenic disturbance in tropical forests can double biodiversity loss from deforestation.*Nature* , *535* , 144–147.

Barnosky, A. D., Hadly, E. A., Bascompte, J., Berlow, E. L., Brown, J. H., Fortelius, M., Getz, W. M., Harte, J., Hastings, A., Marquet, P. A., Martinez, N. D., Mooers, A. O., Roopnarine, P. D., Vermeij, G., Williams, J. W., Gillespie, R. G., Kitzes, J., Marshall, C., Matzke, N., Mindell, D. P., Revilla, E., & Smith, A. B. (2012). Approaching a state shift in Earth's biosphere. *Nature* , *486* , 52–58.

Barnosky, A. D., Matzke, N., Tomiya, S., Wogan, G. O. U., Swartz, B., Quental, T. B., Marshall, C., McGuire, J. L., Lindsey, E. L., Maguire, K. C., Mersey, B., & Ferrer, E. A. (2011). Has the Earth's sixth mass extinction already arrived? *Nature* , *471* , 51–57.

Bohmann, K., Evans, A., Gilbert, M. T. P., Carvalho, G. R., Creer, S., Knapp, M., Yu, D. W., & de Bruyn, M. (2014). Environmental DNA for wildlife biology and biodiversity monitoring. *Trends in Ecology and Evolution* , *29* , 358–367.

Boyer, F., Mercier, C., Bonin, A., Le Bras, Y., Taberlet, P., & Coissac, E. (2016). Obitools: A unix-inspired software package for DNA metabarcoding. *Molecular Ecology Resources* , *16* , 176–182.

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

Cardoso, P., Borges, P. A. V., Triantis, K. A., Ferrández, M. A., & Martín, J. L. (2012). The underrepresentation and misrepresentation of invertebrates in the IUCN Red List. *Biological Conservation* ,*149* , 147–148.

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

Cardoso, P., Silva, I., De Oliveira, N. G., & Serrano, A. R. M. (2004). Indicator taxa of spider (Araneae) diversity and their efficiency in conservation. *Biological Conservation* , *120* , 517–524.

Coddington, J. A,, Young, L. H., & Coyle, F. A. (1996). Estimating Spider Species Richness in a Southern Appalachian Cove Hardwood Forest.*Journal of Arachnology* , *24* , 111–128.

Concato, J., Peduzzi, P., Holford, T. R., & Feinstein, A. R. (1995). Importance of events per independent variable in proportional hazards regression analysis II. *Journal of Clinical Epidemiology* ,*48* , 1503–1510.

Crespo, L. C., Domènech, M., Enguídanos, A., Malumbres-Olarte, J., Moya-Laraño, J., Frías-López, C., Macías-Hernández, N., De Mas, E., Mazzuca, P., Mora, E., Opatová, V., Planas, E., Ribera, C., Ruiz, D., Sousa, P., Tonzo, V., & Arnedo, M. A. (2018). A DNA barcode-assisted annotated checklist of the spider (Arachnida, Araneae) communities associated to white oak woodlands in Spanish National Parks.*Biodiversity Data Journal* , 1–273.

Cribari-Neto, F., & Zeileis, A. (2010). Beta Regression in R.*Handbook of Meta-Analysis* , *34* , 129–150.

Deagle, B. E., Thomas, A. C., McInnes, J. C., Clarke, L. J., Vesterinen, E. J., Clare, E. L., Kartzinel, T. R., & Eveson, J. P. (2019), Counting with DNA in metabarcoding studies: How should we convert sequence reads to dietary data? *Molecular Ecology* , *28* , 391–406.

Dirzo, R., Young, H. S., Galetti, M., Ceballos, G., Isaac, N. J. B., & Collen, B. (2014). Defaunation in the Anthropocene. *Science* ,*345* , 401–406.

Dobyns, J. R. (1997). Effects of Sampling Intensity on the Collection of Spider (Araneae) Species and the Estimation of Species Richness.*Environmental Entomology* , *26* , 150–162.

Eberhard, W. G. (1985). *Sexual Selection and Animal Genitalia* . Harvard University Press, Cambridge, US.

Elbrecht, V., & Leese, F. (2015). Can DNA-based ecosystem assessments quantify species abundance? Testing primer bias and biomass-sequence relationships with an innovative metabarcoding protocol. *PLoS ONE* , *10* , 1–16.

Frøslev, T. G., Kjøller, R., Bruun, H. H., Ejrnæs, R., Brunbjerg, A. K., Pietroni, C., & Hansen, A. J. (2017). Algorithm for post-clustering curation of DNA amplicon data yields reliable biodiversity estimates.*Nature Communications* , *8(1188)* , 1-11.

Garcés-Pastor, S., Wangensteen, O.S., Pérez-Haase, A., Pèlachs, A., Pérez-Obiol, R., Cañellas-Boltà, N., Mariani, S., & Vegas-Vilarrúbia, T. (2019). DNA metabarcoding reveals modern and past eukaryotic communities in a high-mountain peat bog system. *Journal of Paleolimnology* , *62* , 425–441.

Geller, J., Meyer, C., Parker, M., & Hawk, H. (2013). Redesign of PCR primers for mitochondrial cytochrome c oxidase subunit I for marine invertebrates and application in all-taxa biotic surveys.*Molecular Ecology Resources* , *13* , 851–861.

Hebert, P., Ratnasingham, S., & DeWaard, J. R. (2003). Barcoding animal life: Cytochrome c oxidase subunit 1 divergences among closely related species. *Proceedings of the Royal Society B: Biological Sciences,270* , 96-99.

Hoang, D. T., Chernomor, O., von Haeseler, A., Minh, B. Q., & Vinh, L. S. (2018). UFBoot2: Improving the Ultrafast Bootstrap Approximation.*Molecular Biology and Evolution* , *35(2)* , 518–522.

Hulme, P. E. (2009). Trade, transport and trouble: managing invasive species pathways in an era of globalization. *Journal of Applied Ecology* , *46* , 10–18.

Jakob, C., Poizat, G., Veith, M., Seitz, A., & Crivelli, A. J. (2003). Breeding phenology and larval distribution of amphibians in a Mediterranean pond network with unpredictable hydrology.*Hydrobiologia* , *499* , 51–61.

Kennedy, S., Prost, S., Overcast, I., Rominger, A.J., Gillespie, R. G. & Krehenwinkel, H. (2020). High-throughput sequencing for community analysis: the promise of DNA barcoding to uncover diversity, relatedness, abundances and interactions in spider communities.*Development Genes and Evolution* , *230* , 185–201.

Krehenwinkel, H., Wolf, M., Lim, J. Y., Rominger, A. J., Simison, W. B. & Gillespie, R. G. (2017). Estimating and mitigating amplification bias in qualitative and quantitative arthropod metabarcoding.*Scientific Reports* , *7* , 1–12.

Kremen, C., Colwell, R. K., Erwin, T. L., Murphy, D. D., Noss, R. F., & Sanjayan, M. A. (1993). Terrestrial Arthropod Assemblages: Their Use in Conservation Planning. *Conservation Biology* , *7* , 796–808.

Kulikova, T., Aldebert, P., Althorpe, N., Baker, W., Bates, K., Browne, P., van den Broek, A., Cochrane, G., Duggan, K., Eberhardt, R., Faruque, N., Garcia-Pastor, M., Harte, N., Kanz, C., Leinonen, R., Lin, Q., Lombard, V., Lopez, R., Mancuso, R., McHale, M., Nardone, F., Silventoinen, V., Stoehr, P., Stoesser, G., Tuli, M. A., Tzouvara, K., Vaughan, R., Wu, D., Zhu, W., & Apweiler, R. (2004). The EMBL Nucleotide Sequence Database. *Nucleic Acids Research* , *32* , 27–30.

Kuntner, M., & Baxter, I. H. (1997). A preliminary investigation of spider species richness in an eastern Slovenian broadleaf forest.*Proceedings of the 16th European Colloquium of Arachnology* , 173–182.

Lamb, P. D., Hunter, E., Pinnegar, J. K., Creer, S., Davies, R. G., & Taylor, M. I. (2019). How quantitative is metabarcoding: A meta-analytical approach. *Molecular Ecology* , *28* , 420–430.

Lazaridou, M., & Sgardelis, S. (1997). Phenological patterns and life history tactics of Helicoidea (Gastropoda, Pulmonata) snails from Northern Greece. *Iberus* , *15(2)* , 25-34.

Leather, S. R. (2018). "Ecological Armageddon" – more evidence for the drastic decline in insect numbers. *Annals of Applied Biology* ,*172* , 1–3.

Leray, M., Yang, J. Y., Meyer, C. P., Mills, S. C., Agudelo, N., Ranwez, V., Boehm, J. T. & Machida, R. J. (2013). A new versatile primer set targeting a short fragment of the mitochondrial COI region for metabarcoding metazoan diversity: Application for characterizing coral reef fish gut contents. *Frontiers in Zoology, 10* , 1–14.

Levi, H.W. (1974). The orb-weaver genera Araniella and Nuctenea.*Bulletin of the Museum of Comparative Zoology, 146* , 291–316.

Macías-Hernández, N., Ramos, C., Domènech, M., Febles, S., Santos, I., Arnedo, M. A., Borges, P. A. V., Emerson, B. C., & Cardoso, P. (2020). A database of functional traits for spiders from native forests of the Iberian Peninsula and Macaronesia. *Biodiversity Data Journal* ,*8* .

Mahé, F., Rognes, T., Quince, C., de Vargas, C., & Dunthorn, M. (2015). Swarmv2: Highly-scalable and high-resolution amplicon clustering.*PeerJ* , 1–12.

Malumbres-Olarte, J., Boieiro, M., Cardoso, P., Carvalho, R., Crespo, L. C., Gabriel, R., Hernández, N. M., Paulo, O. S., Pereira, F., Rego, C., Ros-Prieto, A., Silva, I., Vieira, A., Rigal, F., & Borges, P. A. V. (2020a). Standardised inventories of spiders (Arachnida, Araneae) of Macaronesia II: The native forests and dry habitats of Madeira archipelago (Madeira and Porto Santo islands). *Biodiversity Data Journal* , *8* .

Malumbres-Olarte, J., Cardoso, P., & Crespo, L. C. (2019). Standardised inventories of spiders (Arachnida , Araneae) of Macaronesia I: The native forests of the Azores (Pico and Terceira islands).*Biodiversity Data Journal* , *7* : e32625.

Malumbres-Olarte, J., Crespo, L. C., Domènech, M., Moya-Laraño, J., & Arnedo, M. A. (2020b). How Iberian are we? Mediterranean climate determines structure and endemicity of spider communities in Iberian oak forests. *Biodiversity and Conservation* , *29* , 3973–3996.

Meiklejohn, K. A., Wallman, J. F., & Dowton, M. (2012). DNA Barcoding Identifies all Immature Life Stages of a Forensically Important Flesh Fly (Diptera: Sarcophagidae). *Journal of Forensic Sciences* ,*58* , 184–187.

Miller, M.A., Pfeiffer, W., & Schwartz, T. (2010). Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In:*Proceedings of the Gateway Computing Environments Workshop (GCE)* , New Orleans, 2010, 1–8.

Minh, B. Q., Nguyen, M. A. T., & von Haeseler, A. (2013). Ultrafast approximation for phylogenetic bootstrap. *Molecular Biology and Evolution* , *305* , 1188–1195.

Morinière, J., Cancian De Araujo, B., Lam, A. W., Hausmann, A., Balke, M., Schmidt, S., Hendrich, L., Doczkal, D., Fartmann, B., Arvidsson, S., & Haszprunar, G. (2016). Species identification in malaise trap samples by DNA barcoding based on NGS technologies and a scoring matrix.*PLoS ONE* , *11* , 1–14.

Nadal, M. F., Achitte-Schmutzler, H. C., Zanone, I., Gonzalez, P. Y., & Avalos, G. (2018). Diversidad estacional de arañas en una reserva natural del Espinal en Corrientes, Argentina. *Caldasia* ,*40* , 129–143.

Nguyen, L.-T., Schmidt, H.A., von Haeseler, A., & Minh, B.Q. (2015) IQ-TREE: A fast and effective stochastic algorithm for estimating maximum likelihood phylogenies. *Molecular Biology and Evolution* ,*32* , 268-274.

Norris, K. C. (1999). Quantifying change through time in spider assemblages: Sampling methods, indices and sources of error.*Journal of Insect Conservation* , *3* , 309–325.

Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., O'hara, R. B., Simpson, G. L., & Wagner H. (2010). Vegan: community ecology package. R package version.

Oliver, I., & Beattie, A. J. (1996). Designing a cost-effective invertebrate survey: A test of methods for rapid assessment of biodiversity. *Ecological Applications* , *6* , 594–607.

Piñol, J., Mir, G., Gomez-Polo, P., & Agustí, N. (2015). Universal and blocking primer mismatches limit the use of high-throughput DNA sequencing for the quantitative metabarcoding of arthropods.*Molecular Ecology Resources* , *15* , 819–830.

R Core Team. (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Rádai, Z., Kiss, B., & Samu, F. (2017). Effect of weather conditions on cohort splitting in a Wolf spider species. *Journal of Arachnology* , *45* , 444–447.

Ratcliffe, F. C., Uren Webster, T. M., Rodriguez-Barreto, D., O'Rorke, R., Garcia de Leaniz, C., & Consuegra, S. (2020). Rapid quantitative assessment of fish larvae community composition using metabarcoding.*bioRxiv* .

Ratnasingham, S., & Hebert, P. (2007). BOLD: The Barcode of Life Data System (http://www.boldsystems.org). *Molecular Ecology Notes* ,*7* , 355–364.

Richard, B., Decaëns, T., Rougerie, R., James, S. W., Porco, D., & Hebert, P. (2010). Re-integrating earthworm juveniles into soil biodiversity studies: Species identification through DNA barcoding.*Molecular Ecology Resources* , *10* , 606–614.

Rognes, T., Flouri, T., Nichols, B., Quince, C., & Mahé, F. (2016). VSEARCH: A versatile open source tool for metagenomics. *PeerJ, 4,*e2584.

Russell-Smith, A., & Stork, N. E. (1995). Composition of spider communities in the canopies of rainforest trees in Borneo. *Journal of Tropical Ecology* , *11* , 223–235.

Sales, N. G., McKenzie, M. B., Drake, J., Harper, L. R., Browett, S. S., Coscia, I., Wangensteen, O. S., Baillie, C., Bryce, E., Dawson, D. A., Ochu, E., Hänfling, B., Lawson Handley, L., Mariani, S., Lambin, X., Sutherland, C., & McDevitt, A. D. (2020). Fishing for mammals: Landscape-level monitoring of terrestrial and semi-aquatic communities using eDNA from riverine systems. *Journal of Applied Ecology* ,*57* , 707–716.

Schenk, J., Geisen, S., Kleinbölting, N., & Traunspurger, W. (2019). Metabarcoding data allow for reliable biomass estimates in the most abundant animals on earth. *Metabarcoding and Metagenomics,3* , 117–126.

Schiebelhut, L. M., Abboud, S. S., Gómez-Daglio, L. E., Swift, H. F., & Dawson, M. N. (2017). A comparison of DNA extraction methods for high-throughput DNA analyses. *Molecular Ecology Resources* ,*17* , 721–729.

Scott, J. A., & Epstein, M. E. (1987). Factors Affecting Phenology in a Temperate Insect Community. *American Midland Naturalist* ,*117* , 103–118.

Silva, D. (1996). Species composition and community structure of Peruvian rainforest spiders: a case study from a seasonally inundated forest along the Samiria river. *Revue suisse de Zoologie* ,*2* , 597–610.

Socolar, J. B., Valderrama Sandoval, E. H., & Wilcove, D. S. (2019). Overlooked biodiversity loss in tropical smallholder agriculture.*Conservation Biology* , *33* , 1338–1349.

Soininen, E. M., Gauthier, G., Bilodeau, F., Berteaux, D., Gielly, L., Taberlet, P., Gussarova, G., Bellemain, E., Hassel, K., Stenøien, H. K., Epp, L., Schrøder-Nielsen, A., Brochmann, C., & Yoccoz, N. G. (2015), Highly overlapping winter diet in two sympatric lemming species revealed by DNA metabarcoding. *PLoS ONE* , *10* , 1–18.

Sørensen, L. L., Coddington, J. A., & Scharff, N. (2002). Inventorying and estimating subcanopy spider diversity using semiquantitative sampling methods in an afromontane forest. *Environmental Entomology* , *31* , 319–330.

Soukainen, A., Pajunen, T., Korhonen, T., Saarinen, J., Chichorro, F., Jalonen, S., Kiljunen, N., Koskivirta, N., Kuurne, J., Leinonen, S., Salonen, T., Yrjölä, V., Fukushima, C., & Cardoso, P. (2020). Standardised spider (Arachnida, Araneae) inventory of Lammi, Finland.*Biodiversity Data* Journal, *8* , 1–15.

Stuart Chapin III, F., Zavaleta, E. S., Eviner, V. T., Naylor, R. L., Vitousek, P. M., Reynolds, H. L., Hooper, D. U., Lavorel, S., Sala, O. E., Hobbie, S. E., Mack, M. C., & Díaz, S. (2000). Consequences of changing biodiversity. *Nature* , *405* , 234–242.

Taylor, R. J., & Doran, N. (2001). Use of terrestrial invertebrates as indicators of the ecological sustainability of forest management under the Montreal Process. *Journal of Insect Conservation* , *5* , 221–231.

Telfer, A. C., Young, M. R., Quinn, J., Perez, K., Sobel, C. N., Sones, J. E., Levesque-Beaudin, V., Derbyshire, R., Fernandez-Triana, J., Rougerie, R., Thevanayagam, A., Boskovic, A., Borisenko, A. V., Cadel, A., Brown, A., Pages, A., Castillo, A. H., Nicolai, A., Mockford, B. M. G., Bukowski, B., Wilson, B., Trojahn, B., Lacroix, C. A., Brimblecombe, C., Hay, C., Ho, C., Steinke, C., Warne, C. P., Cortes, C. G., Engelking, D., Wright, D., Lijtmaer, D. A., Gascoigne, D., Martich, D. H., Morningstar, D., Neumann, D., Steinke, D., DeBruin, D. D. B. M., Dobias, D., Sears, E., Richard, E., Damstra, E., Zakharov, E. V., Laberge, F., Collins, G. E., Blagoev, G. A., Grainge, G., Ansell, G., Meredith, G., Hogg, I., McKeown, J., Topan, J., Bracey, J., Guenther, J., Sills-Gilligan, J., Addesi, J., Persi, J., Layton, K. K. S., D'Souza, K., Dorji, K., Grundy, K., Nghidinwa, K., Ronnenberg, K., Lee, K. M., Xie, L., Lu, L., Penev, L., Gonzalez, M., Rosati, M. E., Kekkonen, M., Kuzmina, M., Iskandar, M., Mutanen, M., Fatahi, M., Pentinsaari, M., Bauman, M., Nikolova, N. I., Ivanova, N. V., Jones, N., Weerasuriya, N., Monkhouse, N., Lavinia, P. D., Jannetta, P., Hanisch, P. E., McMullin, R. T., Flores, R. O., Mouttet, R., Vender, R., Labbee, R. N., Forsyth, R., Lauder, R., Dickson, R., Kroft, R., Miller, S. E., MacDonald, S., Panthi, S., Pedersen, S., Sobek-Swant, S., Naik, S., Lipinskaya, T., Eagalle, T., Decaëns, T., Kosuth, T., Braukmann, T. W. A., Woodcock, T., Roslin, T., Zammit, T., Campbell, V., Dincă, V., Peneva, V., Hebert, P., & DeWaard, J. R. (2015). Biodiversity inventories in high gear: DNA barcoding facilitates a rapid biotic survey of a temperate nature reserve. *Biodiversity Data Journal* ,*3* .

Thomas, A. C., Deagle, B. E., Eveson, J. P., Harsch, C. H., & Trites, A. W. (2016). Quantitative DNA metabarcoding: Improved estimates of species proportional biomass using correction factors derived from control material. *Molecular Ecology Resources* , *16* , 714–726.

Toti, D. S., Coyle, F. A., & Miller, J. A. (2000). A structured inventory of appalachian grass bald and heath bald spider assemblages and a test of species richness estimator performance. *Journal of Arachnology* , *28* , 239–345.

Wangensteen, O.S., Palacín, C., Guardiola, M., & Turon, X. (2018). DNA metabarcoding of littoral hard-bottom communities: High diversity and database gaps revealed by two molecular markers. *PeerJ* , *6* , e4705.

Watts, C., Dopheide, A., Holdaway, R., Davis, C., Wood, J., Thornburrow, D., & Dickie, I. A. (2019). DNA metabarcoding as a tool for invertebrate community monitoring: a case study comparison with conventional techniques. *Austral Entomology* , *58* , 675–686.

Yellareddygari, S. K., Pasche, J. S., Taylor, R. J., Hua, S., & Gudmestad, N. C. (2016). Beta Regression Model for Predicting the Development of Pink Rot in Potato Tubers During Storage. *Plant Disease* , *100* , 1118–1124.

Yen, A., & Butcher, R. (1997). *An overview of the conservation of Non-marine Invertebrates in Australia* . Canberra, Australia, Environment Australia.

Yu, D. W., Ji, Y., Emerson, B. C., Wang, X., Ye, C., Yang, C., & Ding Z. (2012). Biodiversity soup: Metabarcoding of arthropods for rapid biodiversity assessment and biomonitoring. *Methods in Ecology and*

Evolution , *3* , 613–623.

SUPPORTING INFORMATION

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

Table S1 Information on the adult sequenced specimens used in this study.

Table S2 Final dataset containing all the MOTUs of immature stages of spiders obtained by metabarcoding.

Figure S1 Rarefaction curves showing the number of MOTUs per replicate against an increasing number of reads.

FIGURES

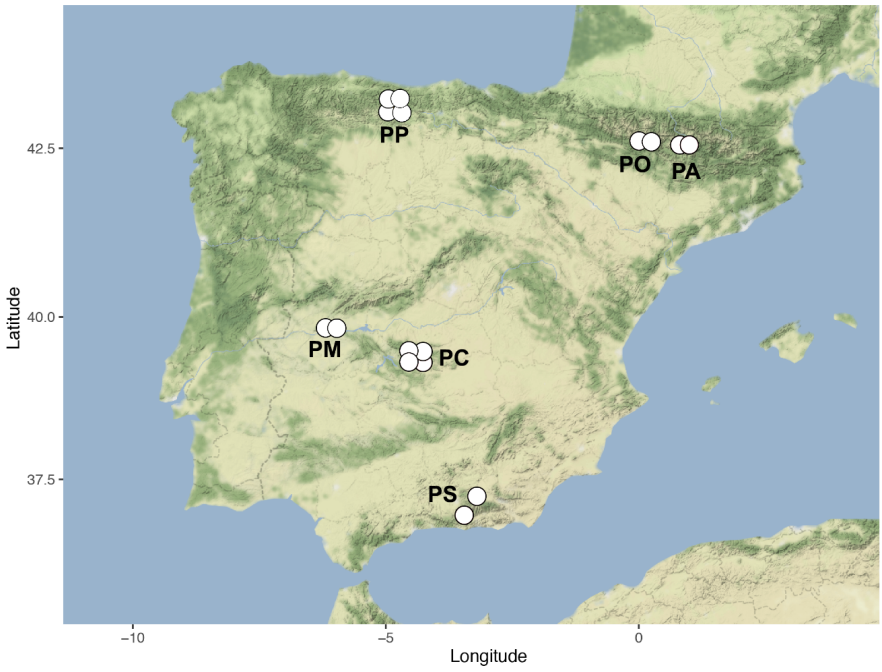

Figure 1. Location of the 16 studied communities. PA – Aigüestortes i Estany de Sant Maurici; PC – Cabañeros; PM – Monfragüe; PO – Ordesa y Monte Perdido; PP – Picos de Europa; PS – Sierra Nevada.

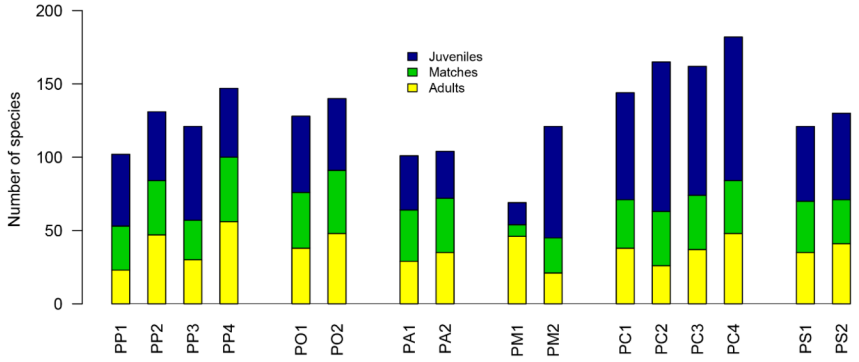

Figure 2. Number of species (MOTUs) found only as adults, only as juveniles and as both life stages in the

16 sampled plots.

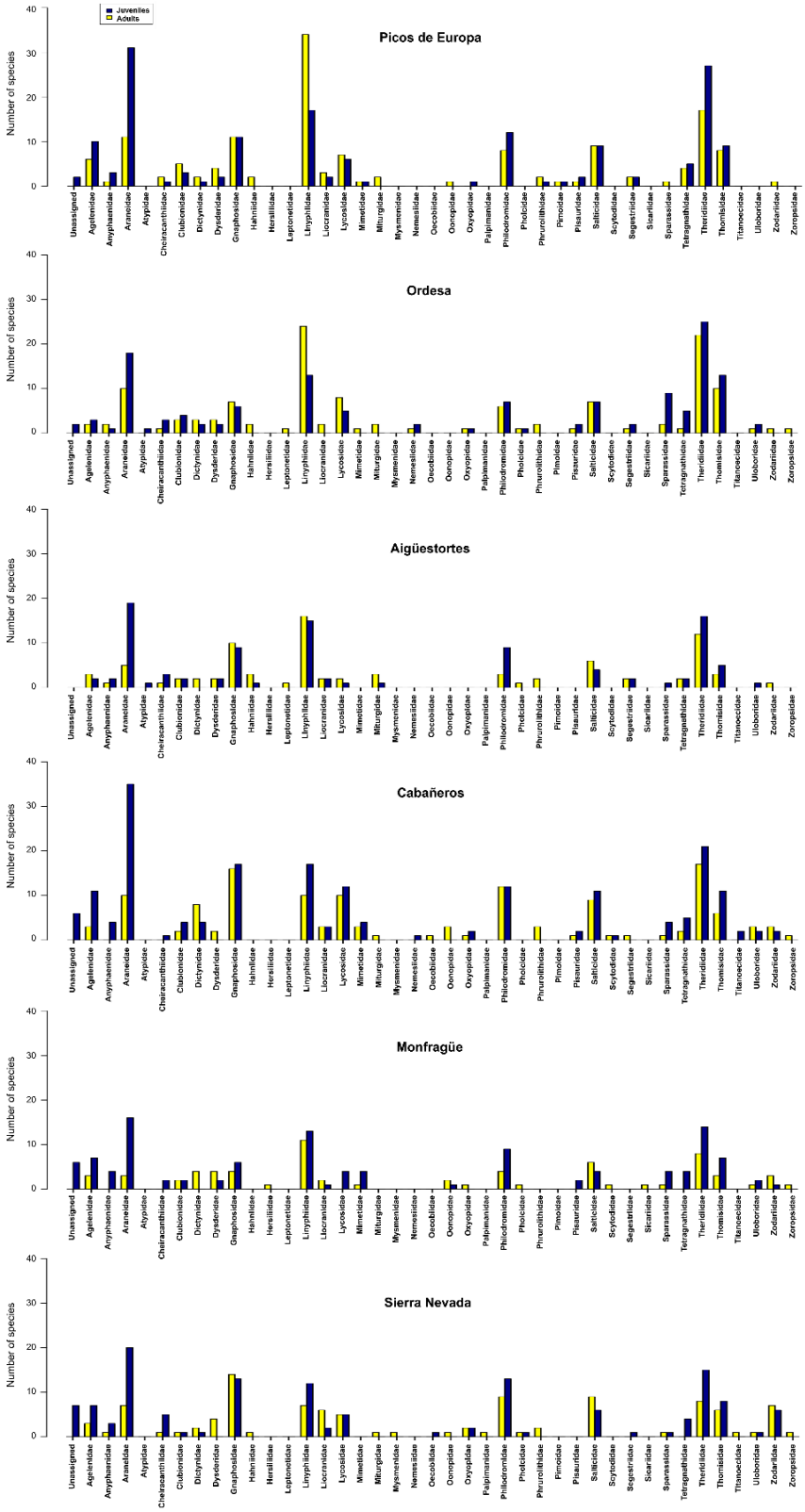

Figure 3. Number of species found as adults (yellow) and number of MOTUs found as juveniles (blue) for each family in every National Park.

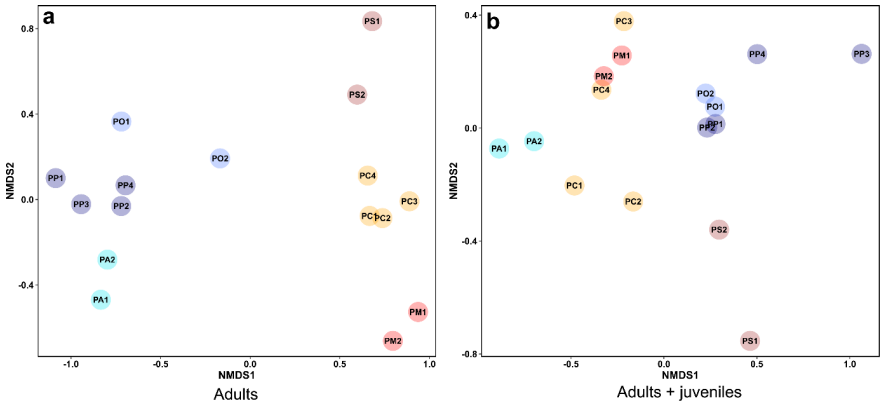

Fig 4. NMDS ordinations of the community composition based on presence/absence data using the Bray-Curtis dissimilarity index. Northern communities are represented in cold colors and southern communities are represented in warm colors. a) NMDS using only adult individuals (stress=0.074). b) NMDS using all individuals (adults and juveniles) (stress=0.065).

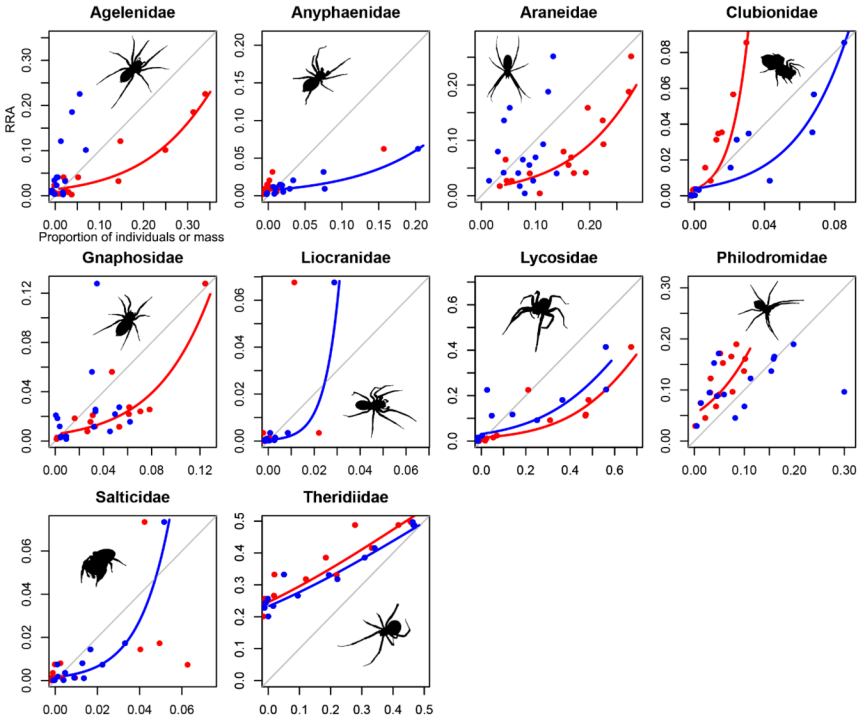

Figure 5. Plots of the relation between relative read abundance (RRA) and proportion of mass (red) and proportion of individuals (blue). Only models with $p > 0.05$ and an adequate goodness of fit are shown.