What DNA barcodes reveal: microhabitat preference, hunting strategy and dispersal ability drive genetic variation across Iberian spider species

MARC DOMÈNECH,[1] JAGOBA MALUMBRES-OLARTE,[2,3]
ALBA ENGUÍDANOS,[1] CESC MÚRRIA[1] and MIQUEL A. ARNEDO[1]   [1]Department
of Evolutionary Biology, Ecology and Environmental Sciences & Biodiversity Research Institute (IRBio), Universitat de Barcelona,
Barcelona, Spain, [2]cE3c – Centre for Ecology, Evolution and Environmental Changes/Azorean Biodiversity Group, Universidade dos
Açores, Angra do Heroísmo, Azores, Portugal and [3]LIBRe – Laboratory for Integrative Biodiversity Research, Finnish Museum of
Natural History, University of Helsinki, Helsinki, Finland

ABSTRACT. 1. The current rate of species loss calls for immediate actions to preserve biodiversity and ecosystem functioning. Cataloguing species richness and composition, and revealing how diversity is geographically distributed are the first steps towards designing efficient conservation strategies.

2. Here, we aim to determine diversity patterns and potential drivers of taxonomic and genetic diversity and population structure of Iberian spiders. We used a community level perspective, analysing more than 3000 DNA barcode sequences representing ∼370 spider species dwelling in white-oak forest habitats across the Spanish National Park network.

3. By combining and comparing morphological and DNA barcode-based species delimitation methods, we assessed their performance and identified putative factors behind cases of incongruence. Our findings uncovered potential overlooked diversity as suggested by the geographic patterns of genetic variation and put a red flag on those taxa that may be undergoing overlooked evolutionary or ecological processes.

4. Spider functional traits associated with foraging strategy, microhabitat preference, ballooning ability and circadian activity explained the observed patterns of population structure across species but did not explain variation in genetic diversity. Overall, our study represents a major step forward in the understanding of large-scale diversity patterns in Iberian spiders at the community level and provides relevant information to guide future conservation strategies of the so-far largely overlooked invertebrate diversity.

Key words. Araneae, community ecology, diversity, DNA barcoding, genetic diversity, population structure.

Introduction

Current biodiversity loss, mostly driven by human activities, is a major concern at the global scale (Barlow *et al.*, 2016; Doherty *et al.*, 2016; Socolar *et al.*, 2019). Species are becoming extinct at a rate a 1000 times faster than the natural background rate of extinction (Pimm *et al.*, 2014).

In this context, understanding diversity patterns and the processes that shape them is of utmost importance for developing more efficient conservation strategies. The first step towards this goal is cataloguing the species present at a given area. However, this may be a daunting task, especially for megadiverse and extremely abundant groups such as arachnids, insects and molluscs (Peters & Wassenberg, 1983; Chapman, 2009), whose morphologic identification is very time-consuming and requires the increasingly more difficult to find specialised taxonomic expertise.

Correspondence: Marc Domènech, Department of Evolutionary Biology, Ecology and Environmental Sciences & Biodiversity Research Institute (IRBio), Universitat de Barcelona, Av. Diagonal 643, E-08028, Barcelona, Spain. E-mail: mdomenan@gmail.com

DNA barcoding approaches are widely used to overcome the limitations of identifying large numbers of specimens (Blagoev *et al.*, 2015; Astrin *et al.*, 2016) by using a small fragment of DNA as a species identifier. The animal DNA barcode is a fragment of around 658 bp from the cytochrome c oxidase subunit I (COI) mitochondrial gene, which has been shown to effectively discriminate between related species (Barrett & Hebert, 2005). DNA barcoding is accelerating diversity inventories by automatising the identification process and provides finer-scale taxonomic resolution than morphological identification. Moreover, DNA barcoding requires little taxonomic expertise when the reference DNA barcode library is already available and provides species names to different life stages or even remnants of individuals (Victor *et al.*, 2009; Zeale *et al.*, 2011). Also, DNA barcode libraries are essential for studies using high-throughput sequencing technologies, such as metabarcoding, which rely entirely on such libraries to assign species names to the obtained sequences (Hajibabaei *et al.*, 2011; Yu *et al.*, 2012).

DNA barcoding has not only been used for species identification but also for preliminary delimitation of species boundaries. Morphological circumscriptions may disagree with molecular-based delimitations. These incongruences may convey different specific ecological and evolutionary processes, such as the presence of cryptic species, deep population structure, introgression/hybridization events, incomplete lineage sorting (ILS), or even misidentifications. Multiple criteria and approaches for delimiting species boundaries based on DNA barcodes or other single markers have been devised (Casiraghi *et al.*, 2010). For instance, the distance-based 'DNA barcode gap' identifies thresholds or cut-off values comparing the differences between the lowest interspecific genetic distance and the highest intraspecific distance by conducting pairwise comparisons among all available sequences. Other more sophisticated species delimitation methods include the Barcode Index Number (BIN) system, which is an example of a genetic distance algorithms method implemented in the Barcode of Life Data System (BOLD) (https://www.boldsystems.org/index.php), whereas the Poisson

Tree Processes method (PTP) (Zhang *et al.*, 2013) and its multi-rate version mPTP (Kapli *et al.*, 2017) are phylogeny-aware approaches.

Intraspecific genetic diversity is a measure of the overall heritable variation within a species. Estimates of intraspecific genetic diversity are affected by demographic fluctuations in effective population size, mutation rates and coalescence time (Ellegren & Galtier, 2016; Salinas-Ivanenko & Múrria, 2021). For instance, habitat stability favours large effective population sizes and thus increases genetic diversity (Papadopoulou *et al.*, 2009; Múrria *et al.*, 2015). The population structure describes how intraspecific genetic diversity is spatially distributed within and across populations. Several factors can affect population structuring, such as population bottlenecks, life history, selection or gene flow (Charlesworth & Willis, 2009; Slatkin, 1987). Isolation by distance (IBD) analyses are commonly used to elucidate the spatial patterns of genetic variation (Schäfer *et al.*, 2001; Pope *et al.*, 2006). IBD is found when the relationship between genetic differentiation among populations and geographic distance is positive. Alternative patterns indicate that different underlying processes are acting on the group of study. For example, high levels of population genetic differentiation regardless of the geographic distance may uncover overlooked species, while the lack of genetic differentiation across large areas may indicate high levels of genetic flow as a result of high connectivity or dispersal ability (Fig. 1).

Spiders are among the most diverse orders of arthropods with 128 families and more than 49 000 species described (World Spider Catalog, 2020). Spiders play a key role as top predators of arthropods in most terrestrial ecosystems (Nyffeler & Birkhofer, 2017) and are in turn a valuable component of the diet of other predators such as birds (Ramsay & Houston, 2003). Because of their species-richness and abundance, morphological identification of spiders is a daunting task. Even more challenging, immature stages are usually not identifiable due to the lack of developed genitalic structures, which are the main source of species diagnostic characters. In this context, the use of DNA
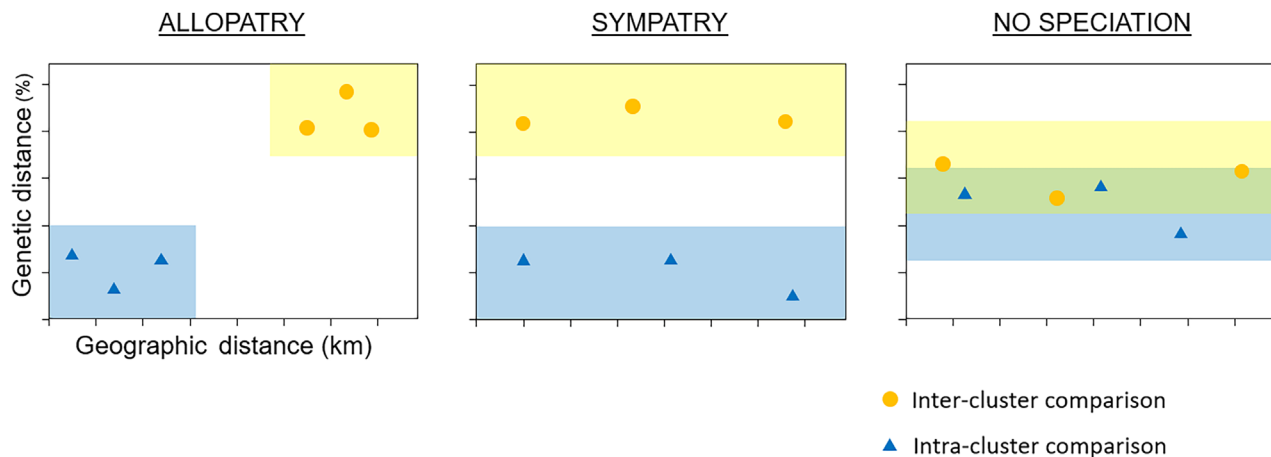

Fig 1. IBD plots representing three possible scenarios (allopatry, sympatry and no lineage differentiation) based on the geographic patterns of intra- and inter-cluster pairwise genetic divergence.

barcoding approaches for accelerating spider inventories has become increasingly popular (e.g. Astrin *et al.*, 2016; Tyagi *et al.*, 2019; Kennedy *et al.*, 2020).

Spiders provide a diversity of measurable functional traits, which may impact their genetic diversity and population structure. While some spiders are very poor dispersers and have restricted ranges, others are capable of long-distance dispersal. For instance, ballooning is a spider-specific strategy of passive aerial dispersal consisting in the use of a thread of silk as a parachute, which is then carried by wind currents and atmospheric electric fields (Morley & Robert, 2018). Several studies found ballooning dispersal behaviour to be related to small body sizes (Dean & Sterling, 1985; Greenstone *et al.*, 1987) and vegetation structure (Blandenier, 2009). Dispersal by ballooning is associated with high gene flow and weak population structure, even in mountainous landscapes (Botham *et al.*, 2020) and can influence the occurrence and the composition of spider communities (Bonte *et al.*, 2003a, 2004; Jiménez-Valverde *et al.*, 2010). In contrast, the level of habitat specialisation seems to be negatively related to the ballooning ability (Bonte *et al.*, 2003b), as a high tendency to balloon increases the risk of reaching potential unsuitable habitats and therefore should increase population structuring. Habitat specialisation has also been found to have a weak or no contribution to explain variation in distribution patterns (Bonte *et al.*, 2004).

Although several studies have investigated genetic diversity and population structure in spiders for describing phylogeographic processes (Ramírez & Chi, 2004; Settepani *et al.*, 2014), they mostly focused on a single or few species. Conversely, phylogenetic studies usually consider large numbers of species, but the intraspecific analyses are either weak or absent. The novelty of our study is to relate the geographic distribution of intraspecific genetic variation to functional traits using a multi-species approach. Therefore, our purpose is not to taxonomically redefine the species studied, but to put a red flag on those groups that may be undergoing interesting evolutionary or ecological processes and whose taxonomy would be worth revising. Here, we aim to determine the diversity patterns at the community level and the drivers of taxonomic and genetic diversity of Iberian spiders. The specific objectives are (i) to build a DNA barcode reference library of Iberian spiders to facilitate and automatise spider identification; (ii) to assess the ability of molecular tools used at the community level to delimit Iberian spiders species compared to morphology, and to investigate cases of incongruence; (iii) to uncover morphologically overlooked diversity (i.e. cryptic diversity) assessing geographic patterns of genetic variation; and (iv) to determine relations between spider functional traits and patterns of intraspecific nucleotide genetic diversity and population structure across species.

Materials and methods

Samplings and study area

We collected the studied spiders in May–June 2013 and 2014. The sampling design consisted in 16 1 -ha plots distributed in white oak forests in six national parks across the Iberian Peninsula (Fig. 2). We used the semi-quantitative COBRA (Conservation Oriented Biodiversity Rapid Assessment) sampling protocol (Cardoso, 2009) for capturing spiders. COBRA combines timed direct capture, foliage beating, vegetation sweeping and pitfall trapping (active for 2 weeks) (see details in the study by Crespo *et al.*, 2018).

DNA extraction, amplification and sequencing

We stored the spiders in absolute ethanol at $-20\,^{\circ}$C, and we used a ZEISS Stemi 2000 stereomicroscope and the 'Araneae: Spiders of Europe' database to identify most of the species (Nentwig *et al.*, 2017). We performed individual DNA extractions and amplified a 658 bp fragment of the animal DNA barcode, the mitochondrial gene cytochrome c oxidase subunit I (COI), by PCR as detailed in the study by Crespo *et al.* (2018), using LCO1490 and HCO2198 (Folmer *et al.*, 1994) as the referred combination of primers. PCR products were cycle-sequenced in both directions at Macrogen Inc. (Spain). There was no evidence of nuclear copies of mitochondrial DNA (NUMTs), and in the cases with presence of non-target sequences (proteobacteria or contamination), we repeated the amplification and sequencing. We uploaded all sequences to BOLD and Genbank (Bensons *et al.*, 2017) (accession numbers MH630465-MH630994, MT215600-MT215716, MT607653-MT607961, MW996898-MW999147 and MZ824404).

Species delimitation analyses

To assess the ability of our DNA barcode library to unequivocally assign a sequence to a particular species, we performed an initial barcode gap analysis for every species by comparing the highest intraspecific uncorrected *P*-distance to the lowest interspecific (nearest neighbour) distance using MEGA7 v7.0.26 (Kumar *et al.*, 2016). The existence of a barcode gap, i.e., a separate distribution of the frequency of intraspecific and interspecific distances, guarantees that unknown sequences can be readily assigned to a species. Alternatively, a small or absent barcode gap blurs species boundaries, making it difficult or impossible to associate a given sequence to a unique species. Albeit fast, the DNA barcode gap approach may suffer from inaccuracy or lead to spurious results (Spooner, 2009; Elias *et al.*, 2009; Hamilton *et al.*, 2014). *Xysticus ferrugineus* Menge, 1876, *Alopecosa cuneata* (Clerck, 1757) and *Phrurolithus nigrinus* (Simon, 1878) contained only one individual so their barcode gap could not be computed.

We also performed molecular species delimitation analyses using two contrasting methods, namely the graph-theory-based BIN system and the character-based mPTP model. BIN follows a Refined Single Linkage (RESL) algorithm combined with Markov clustering. Taking into account all records in the BOLD database, all sequences >500 bp were automatically assigned to BIN clusters, which included highly similar sequences and represent molecularly delimited candidate species. Alternatively, new BINs were created if a good match was not found (Ratnasingham & Hebert, 2013). The mPTP delimitation relies on the

relative branch length information provided by a gene tree to identify a transition point between inter- and intra-specific divergences, i.e., distinguishing putative speciation from coalescence events. We inferred a maximum-likelihood gene tree from all the COI sequences. We aligned them using the online version of MAFFT v7 (Katoh *et al.*, 2019) with the G-ins-I algorithm. We used the IQ-TREE software v1.6.1 (Nguyen *et al.*, 2015; Chernomor *et al.*, 2016) to infer a maximum likelihood tree using the edge-linked partition model and the GTR + I + G model of nucleotide substitution, as it is the most complex and parameter-rich model and the large number of sequences allowed for a correct estimation of the parameters. The analysis was partitioned by codon position. We assessed branch support using ultrafast bootstrap with 1000 replicates. We ran phylogenetic analyses remotely at the CIPRES cluster (Miller *et al.*, 2010), and we ran the mPTP model locally using the programme mPTP v 0.2.4 (Kapli *et al.*, 2017, available at https://github.com/Pas-Kapli/mptp). We first identified the minimum branch length (0.000888), which was subsequently used to delimit clusters and assess their support by means of three Markov Chain Monte Carlo sampling chains ran for 50 million generations, discarding the first 2 million MCMC steps as burnin.

We compared the morphologically defined species to the results of the two molecular delimitation methods. We classified molecular clusters in 'match', 'split', 'merge' and 'mixture' depending on whether they assigned specimens to the same morphological species, split specimens of the same species into more than one cluster, included all the specimens of more than one species, or partially combined specimens from different species, respectively.

IBD analyses

We assessed patterns of IBD for all species involved in 'splits' in both BIN and mPTP by plotting the uncorrected pairwise genetic distance between pairs of individuals against their geographic distance. For this analysis, we assigned unassigned specimens due to their short sequence length to their closest BIN based on the topology of the COI tree. This was the case of *Eratigena picta* (Simon, 1870) MD321, which was assigned to the BIN BOLD:ADM0499, *Ero aphana* (Walckenaer, 1802) MD323 to BOLD:AAO2255, *Paidiscura pallens* (Blackwall, 1834) MD1405 to BOLD:AAO0598, *Paidiscura pallens* MD1410 to BOLD:AAO0598, *Zodarion styliferum* (Simon, 1870) MD3518 to BOLD:ADM0571 and *Zodarion styliferum* MD3519 to BOLD:ADM0571.

We used IBD to assess whether the groups resulting from cases of molecular split of morphologically delimited species were randomly distributed across sites, reflected population structure within a species or, alternatively, supported the existence of co-occurring independent evolutionary lineages (incipient species). To do so, we assigned each morphospecies to 'allopatric', 'sympatric' or 'no-speciation' split categories based
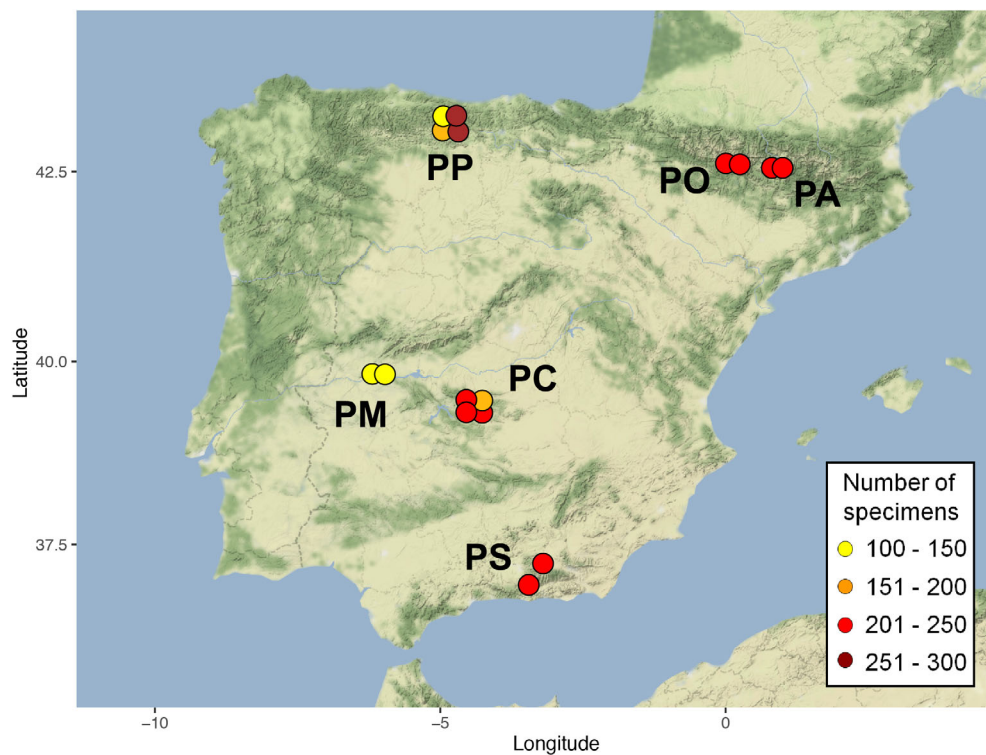

Fig 2. Collection sites for the 3207 spider specimens included in this study. Colours indicate the number of specimens sequenced from each locality (PP: Picos de Europa; PO: Ordesa y Monte Perdido; PA: Aigüestortes i Estany de Sant Maurici; PM: Monfragüe; PC: Cabañeros; PS: Sierra Nevada).

on the patterns recovered in their IBD plots (see Fig. 1). In the cases of allopatric or sympatric splits, specimens were revised in search for possibly overlooked morphological differences.

Genetic diversity, population structure and functional traits

We ran the analyses of intraspecific genetic diversity and population structure only for species with five or more individuals captured in at least two parks (Table 1). For the intraspecific genetic diversity, we calculated the nucleotide diversity (π) of each species as the average number of nucleotide differences per site with the *nuc.div* function in the package 'pegas' (Paradis, 2010) in R v.3.6.3 (R Core Team, 2020). We assessed the level of genetic structure among populations (Phi$_{PT}$) by means of an Analysis of MOlecular VAriance (AMOVA) performed also with 'pegas'. We used the package 'ape' (Paradis & Schliep, 2018) to compute the input matrix of Euclidean genetic distances for the AMOVA analysis, setting the model to *raw* and

pairwise deletion to *false* and we assessed the significance by performing 10 000 permutations. Negative AMOVA values (intra-population differentiation higher than inter-population differentiation) can have different interpretations, but they are usually the result of artefacts due to relatively small sample sizes and we treated them as 0 (Meirmans & Hedrick, 2011).

We investigate the relationship between intraspecific genetic diversity and nine functional traits. These traits included four morphological measurements, namely (i) width and (ii) height of the prosoma and (iii) length of the fang and (iv) of the first tibia. We made the latter two proportional to prosoma width to avoid correlation, and we standardised all morphological traits by substracting the mean and dividing by standard deviation. The remaning traits corresponded to five ecological traits, namely (i) the dispersal ability by ballooning (three categories), (ii) circadian activity (diurnal or nocturnal), (iii) foraging strategy (seven categories), (iv) trophic specialisation (generalist or specialist) and (v) preferred microhabitat or vegetation layer as inferred by the capture methods (four categories from ground

Table 1. List of the 67 species used for population structure and genetic diversity analyses with the number of individuals included from different sites (and the total).

Species	Individuals	Species	Individuals
Dipoena melanogaster	10, 10, 10, 20, 21 (71)	*Lathys humilis*	10, 10 (20)
Theridion harmsi	6, 6, 9, 9, 10, 19 (59)	*Araniella opisthographa*	9, 10 (19)
Eratigena picta	6, 6, 10, 10, 17 (49)	*Pardosa hortensis*	5, 14 (19)
Mangora acalypha	5, 7, 11, 20 (43)	*Salticus scenicus*	7, 12 (19)
Heliophanus cupreus	5, 5, 9, 10, 13 (42)	*Textrix denticulata*	9, 10 (19)
Philodromus fuscolimbatus	7, 34 (41)	*Cyclosa algerica*	8, 9 (17)
Eratigena montigena	10, 11, 18 (39)	*Micaria fulgens*	7, 10 (17)
Philodromus dispar	8, 10, 21 (39)	*Selamia reticulata*	7, 10 (17)
Platnickina tincta	10, 10, 19 (39)	*Ballus chalybeius*	6, 10 (16)
Paidiscura pallens	7, 11, 20 (38)	*Ero aphana*	8, 8 (16)
Pardosa lugubris	5, 9, 22 (36)	*Haplodrassus signifer*	5, 11 (16)
Drassodes lapidosus	10, 10, 14 (34)	*Heterotheridion nigrovariegatum*	6, 10 (16)
Oxyopes nigripalpis	10, 22 (32)	*Nigma puella*	5, 11 (16)
Lasaeola tristis	6, 7, 18 (31)	*Tenuiphantes tenuis*	5, 11 (16)
Phrurolithus szilyi	6, 8, 17 (31)	*Lathys* Dictynidae01	5, 10 (15)
Brigittea civica	5, 9, 16 (30)	*Liocranum rupicola*	6, 9 (15)
Zilla diodia	6, 10, 14 (30)	*Microneta viaria*	5, 5, 5 (15)
Brigittea latens	8, 10, 11 (29)	*Theridion varians*	6, 9 (15)
Clubiona comta	5, 7, 16 (28)	*Tmarus staintoni*	6, 9 (15)
Haplodrassus silvestris	8, 20 (28)	*Agroeca inopina*	7, 7 (14)
Philodromus lividus	10, 18 (28)	*Alopecosa albofasciata*	5, 9 (14)
Theridion mystaceum	7, 20 (27)	*Mimetus laevigatus*	7, 7 (14)
Cyclosa conica	10, 16 (26)	*Diaea dorsata*	5, 8 (13)
Parasteatoda lunata	10, 16 (26)	*Evarcha falcata*	5, 8 (13)
Araniella cucurbitina	5, 7, 12 (24)	*Gibbaranea gibbosa*	5, 8 (13)
Frontinellina frutetorum	5, 8, 10 (23)	*Episinus algiricus*	5, 7 (12)
Micaria guttigera	11, 12 (23)	*Macaroeris nidicolens*	5, 7 (12)
Zoropsis media	5, 8, 10 (23)	*Monocephalus fuscipes*	6, 6 (12)
Civizelotes civicus	5, 17 (22)	*Scytodes velutina*	5, 7 (12)
Clubiona brevipes	5, 7, 10 (22)	*Callilepis concolor*	5, 6 (11)
Zora manicata	5, 7, 10 (22)	*Philodromus albidus*	5, 6 (11)
Drassyllus praeficus	10, 11 (21)	*Tenuiphantes flavipes*	5, 6 (11)
Civizelotes medianus	10, 10 (20)	*Drassyllus villicus*	5, 5 (10)
Eratigena fuesslini	10, 10 (20)		

Species are sorted by total abundance. A more detailed table can be found in supplementary material (Table S1).
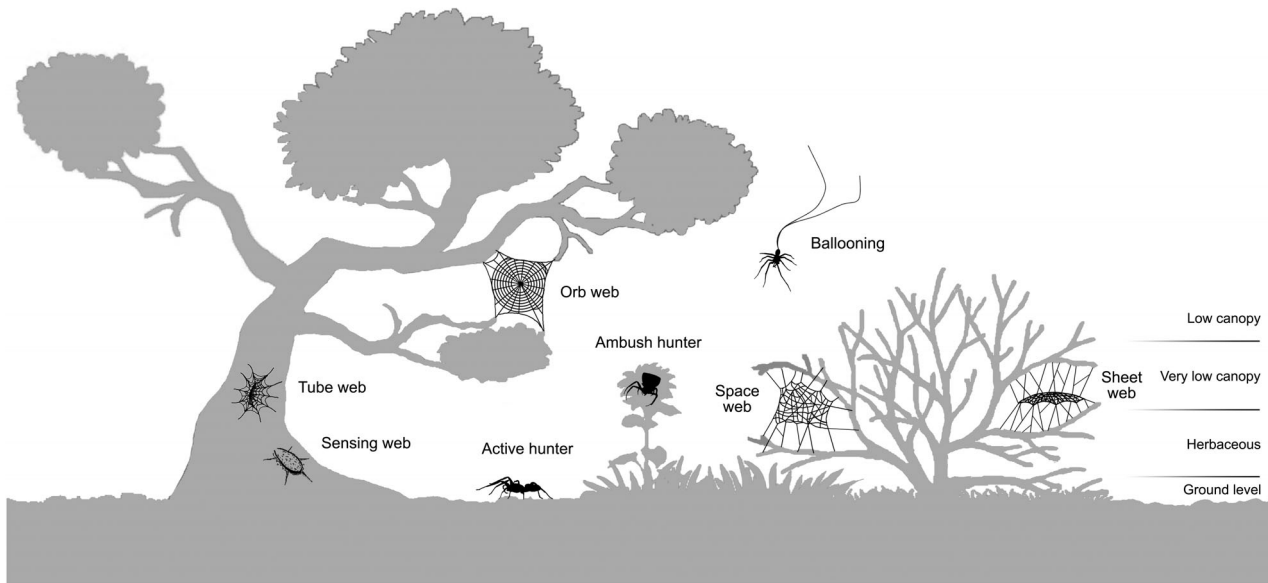
Fig 3. Graphic summary of the ecological traits. Silhouettes depict some traits used in this study, namely dispersal by ballooning, preferred microhabitat or vegetation layer and foraging strategy. Foraging strategy includes five types of web (orb webs, space webs, sheet webs, tube webs and sensing webs such as those of trap-door species) and two types of non-web hunting strategies (ambush hunters and active hunters). Spiders and webs are not to scale.

level to low canopy) (Fig. 3) (see Table S1 for trait codification, and see Macías-Hernández *et al.*, 2020b for trait relevance and justification). We used the R package '*ade4*' (Dray & Dufour, 2007) to generate a Gower dissimilarity matrix to which we applied a principal coordinates analysis (PCoA) to represent the overall functional space including all species. We controlled for the effect of phylogenetic relatedness in statistical models using the time-calibrated phylogeny in the study by Macías-Hernández *et al.* (2020a). This phylogeny was inferred from six genes and topological constraints based on transcriptome information (Fernandez *et al.*, 2018) for 1450 species of spiders, including those in the present study. We pruned the phylogenetic tree retaining only our species of interest using the function *drop.tip* in the package '*ape*'. In the PCoA, we tested the phylogenetic trait conservatism with Pagel's lambda (λ) (Pagel, 1999) using the position of each species on the PCo1 and PCo2 axes with the *phylosig* function in the R package '*phytools*' (Revell, 2012). Values of λ close to 1 indicate that there is a strong phylogenetic signal, while values close to 0 indicate a low phylogenetic signal (i.e. trait states are more randomly distributed across the phylogeny). We assessed the Spearman correlation of each trait on the PCo1 and PCo2 axes using the *cor* function in the R package '*stats*' (R Core Team, 2020).

We ran phylogenetic generalised least squares models (PGLS) using the function *pgls* in the R package '*Caper*' (Orme *et al.*, 2013) to test the correlation between population structure Phi$_{PT}$ (obtained from AMOVA analyses) and the position on the PCo1 and PCo2 axes, as well as between the nucleotide diversity

π and the position on the PCo1 and PCo2 axes, by taking into account the phylogenetic non-independence between species.

Results

We successfully sequenced 3207 barcodes corresponding to 1344 distinct haplotypes. They belonged to 371 morphospecies out of the 377 (98.4%) previously identified using morphology (see Crespo *et al.*, 2018 for details) and to 39 families (70% of those reported in the Iberian Peninsula). The morphospecies in the dataset were represented by nine individuals on average and 75 of them (20.2%) were singletons. Most haplotypes were exclusive of a single individual, but 499 haplotypes were recovered in more than one individual. The number of individuals sharing the same haplotype ranged from 2 to 47 (in *Theridion harmsi* Wunderlich, 2011). Most species showed a clear barcode gap (Supporting Information Fig. S1), insofar as the average maximum intraspecific divergence was 2.07%, while average distance to the nearest neighbour was 9.73%.

BIN analysis

Only 95 COI sequences could not be assigned to any BIN because they did not comply with the minimum length requirements, which left 13 species and four families out of this analysis. The BIN system clustered the remaining

sequences into 441 BINs, in contrast with the 358 identified morphospecies. The correlation between the number of BINs for a species and the number of specimens analysed was significant, but weak ($r_s = 0.1$, $P < 0.001$). The BIN/morphospecies ratio for the 35 families ranged between 1 and 2, with an average of 1.3. There was a perfect match between 269 morphospecies (75.1%) and a particular BIN, while 70 morphospecies (19.6%) were split into more than

one BIN. The average number of BINs per species in 'split' cases was 2.1, ranging from 2 to 6. The cases of 'merge' were uncommon, affecting only 16 morphospecies (4.5%) and in all the cases, they involved pairs of species. There was one single case (0.8%) of 'mixture', which involved specimens of three closely related species of the *Theridion melanurum* Hahn, 1831 species complex, which shared six BINs.
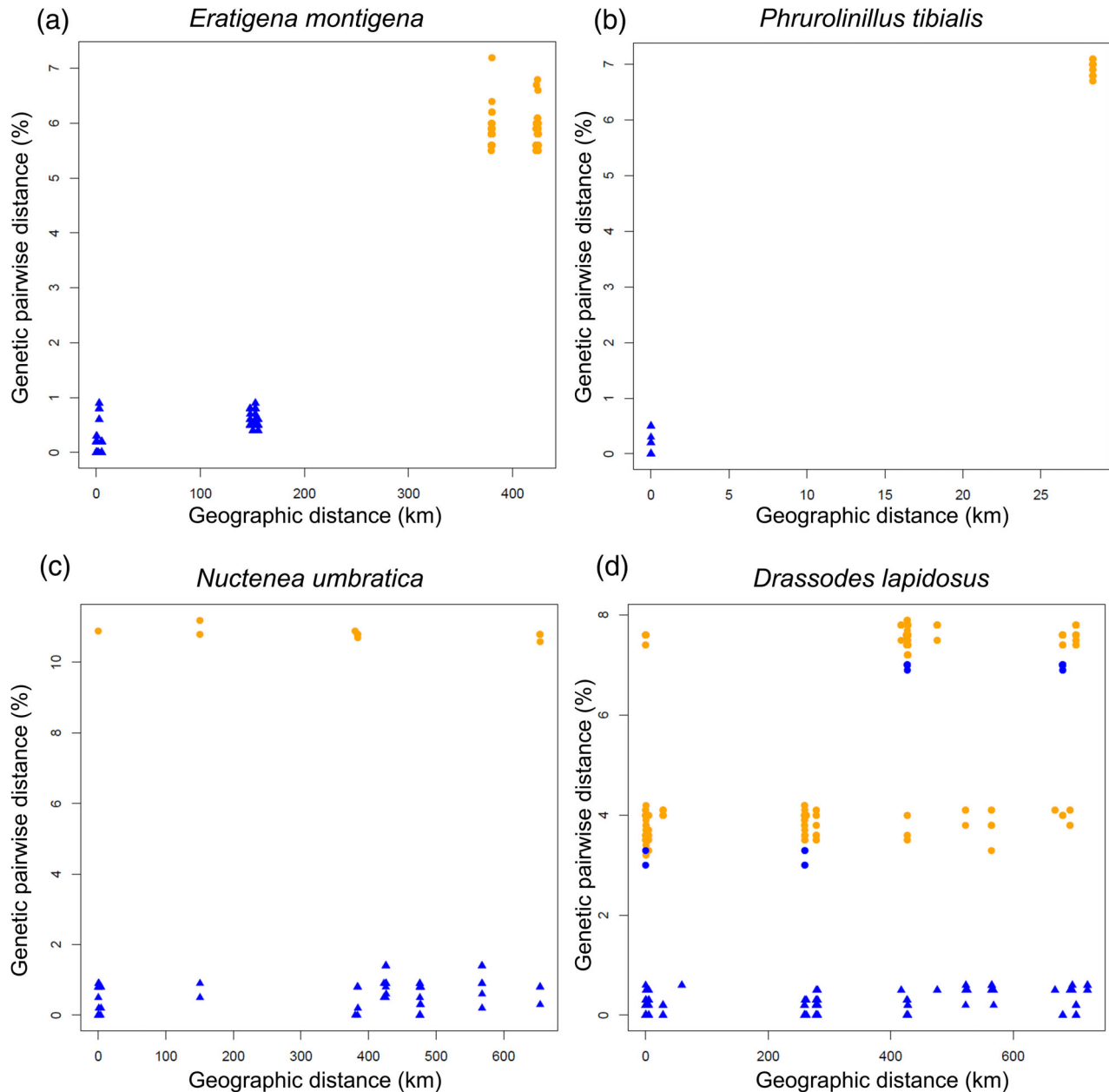

Fig 4. Isolation by distance (IBD) plots of the genetic distance (%) in relation to the geographic distance (km) of representative morphospecies split in both delimitation methods. The plots illustrate cases of allopatric patterns (a and b), sympatric patterns (c) and no apparent geographic pattern (d). Intra- and inter-cluster distances according to BIN delimitation are represented with triangles and circles, respectively. Intra- and inter-cluster distances according to mPTP delimitation are represented in blue and orange, respectively.

mPTP analysis

The inferred phylogenetic tree (Fig. S2) was well resolved, with supports being in general higher towards the tips. For the 34 families represented by more than one taxon, 24 were recovered as monophyletic. At the morphospecies level, all except 16 were monophyletic. The mPTP model sorted the 3207 sequences (371 morphospecies) into 365 clusters. Again, the correlation between the number of clusters and the number of sampled specimens per species was significant, but weak ($r_s = 0.07$, $P < 0.001$). The ratio between the number of clusters and morphospecies for the 39 families ranged from 0.25 to 1.71, with an average of 0.9. Cases of 'match' represented 66.6% of the morphospecies, whereas 'merges' (24%) and 'splits' (8.9%) were less common, and 'mixtures' were marginal (0.5%). The number of cases of 'merge' for the mPTP model were higher than in BIN analyses. These 'merge' cases were found in 26 pairs, seven triads, two tetrads and one pentad. However, 15 of the 'merge' cases included morphospecies from different genera, and three of them were even from different families, which hints to some sort of artefact. The mPTP model found fewer 'splits' than the BIN delimitation, affecting 33 morphospecies that were divided in two or more clusters. The average number of clusters per species in cases of 'split' was 2.3, ranging from 2 to 6. The mPTP model found one single case of 'mixture', which surprisingly involved the distantly related linyphiids *Microneta viaria* (Blackwall, 1841) and *Neriene clathrata* (Sundevall, 1830).

Barcode gap

The analysis of maximum intraspecific vs. minimum interspecific (nearest neighbour) distances revealed the existence of a barcode gap in most morphospecies (94%), which indicates a clear cut-off value between intra- and inter-specific genetic distance. These morphospecies are plotted above the diagonal line in the Supporting Information Fig. S1. Those morphospecies consistently merged both in BIN and mPTP delimitations, showed low intraspecific and interspecific divergences and appeared at the bottom left end of the barcode gap plot, close to the diagonal line (grey triangles in the Supporting Information Fig. S1). Among those morphospecies, nearest neighbour distance was still greater than maximum intraspecific distance in *Civizelotes civicus* (Simon, 1878) (0.8% > 0.7%), *Phrurolithus minimus* C. L. Koch, 1839 (0.8% > 0%) and *Phycosoma inornatum* (O. Pickard-Cambridge, 1861) (1.7% > 0.2%), fulling the requirement for a barcode gap. Conversely, no barcode gap was observed, this is the nearest neighbour distance was smaller than maximum intraspecific distance, in *Agyneta pseudorurestris* Wunderlich, 1980 (0% < 0.3%), *Agyneta rurestris* (C. L. Koch, 1836) (0% < 0.7%), *Alopecosa pulverulenta* (Clerck, 1757) (1.2% < 1.5%), *Civizelotes medianoides* (Senglet, 2012) (0.8% < 1%), *Xysticus audax* (Schrank, 1803) (0.2% < 2%) and *Xysticus cristatus* (Clerck, 1757) (0.2% < 1.2%). For *Dipoena umbratilis* (Simon, 1873) and *Xysticus* sp. 42 both distances were identical (1.7% and 1.4%, respectively). The morphospecies pairs *Agyneta rurestris-A. pseudorurestris* and *Theridion mystaceum* L. Koch, 1870-*T. promiscuum*

(Domènech *et al.*, 2020) both shared haplotypes. Finally, the morphospecies *Zodarion styliferum*, which was split by both delimitation methods, showed the most extreme values in the barcode gap plot, since its nearest neighbour distance (5.8%) was fourfold lower than its maximum intraspecific distance (21.2%). Surprisingly, two species, namely *Haplodrassus silvestris* (Blackwall, 1833) and *Monocephalus fuscipes* (Blackwall, 1836), contained one single haplotype each despite having populations in distant parks.

IBD analyses

The total number of morphospecies consistently involved in 'split' cases in both BIN and mPTP methods was 27. In general, all these morphospecies were split in the exact same way by both delimitation methods, except for: *Aelurillus luctuosus* (Lucas, 1846), *Drassodes lapidosus* (Walckenaer, 1802), *Eratigena picta*, *Evarcha falcata* (Clerck, 1757) and *Segestria senoculata* (Linnaeus, 1758). Eleven out of the 27 morphospecies showed an 'allopatric' pattern (Fig. 1), whereas 13 species showed a 'sympatric' pattern. The remaining three species did not conform to any of the former patterns. Figure 4 illustrates the IBD plots of four specific cases (see Supplementary Material Fig S3 for the IBD plots of all 27 species). *Eratigena montigena* (Fig. 4a) and *Phrurolinillus tibialis* (Fig. 4b) were examples of geographic divergence and allopatric splits, with different genetic clusters corresponding to different localities. For *E. montigena*, the pairwise distances between distant localities (5.5–7.2%) were above the average maximum intraspecific divergence (2.07%), while they were low within the same or nearby localities (0–0.9%) and corresponded to intracluster distances. Similarly, the pairwise distances of *P. tibialis* within each plot ranged from 0%
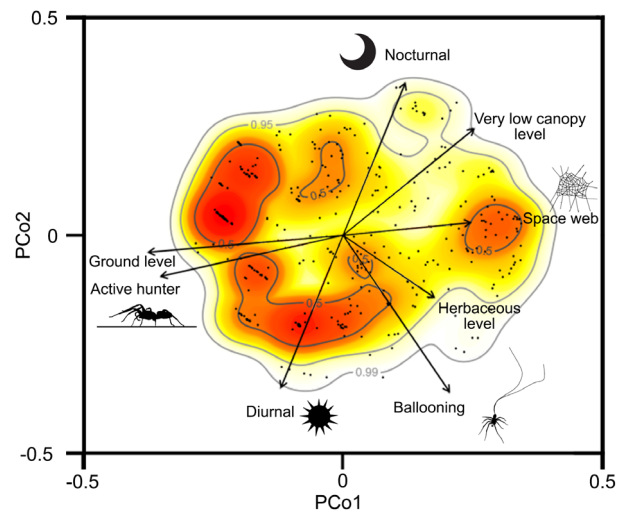

Fig 5. Principal coordinates analysis showing the position of each species (dots) in the functional space. Red colour indicates areas with higher density of species. Arrows indicate the direction of the traits that were correlated with either of the principal coordinates. The length of the arrows was halved for visualisation purposes.

to 0.5%, while the distances between its two populations, barely separated by 28 km, were much higher (6.7–7.1%). The species *Nuctenea umbratica* exemplified a pattern of 'sympatric' divergence (Fig 4c), with low within-cluster (0–1.4%) but high between-cluster pairwise divergences (10.6–11.2%) found within the same locality. Finally, *Drassodes lapidosus* provided an example of the absence of a clear pattern in the distribution of genetic variability (Fig 4d).

Links between patterns of genetic diversity and population structure and functional traits

The first and second PCo-axes of the functional space accounted for 3.1% and 2.3% of the total variance, respectively (Fig. 5). The PCo1 was positively correlated with the 'space web' foraging strategy ($r^2 = 0.5$) and microhabitats 'herbaceous' ($r^2 = 0.51$) and 'very low canopy' ($r^2 = 0.61$) and negatively correlated with the 'active hunter' foraging strategy ($r^2 = -0.72$) and the 'ground level' microhabitat ($r^2 = -0.79$). PCo2 was positively correlated with nocturnal activity ($r^2 = 0.7$) and negatively correlated with diurnal activity ($r^2 = -0.7$) and with the dispersal ability by ballooning ($r^2 = -0.74$) (Table S2). A strong phylogenetic signal was found for both PCo1 ($\lambda = 1.02$, $P < 0.001$) and PCo2 ($\lambda = 0.97$, $P < 0.001$).

For 67 species, at least five individuals were distributed in at least two parks. The unidentified morphospecies *Lathys* sp.1 lacked some functional trait information and was removed from further analyses. The remaining 66 species were used for testing the relationship between functional traits and intraspecific genetic variation. PGLS models found a significant negative correlation between genetic population structure (Phi_{PT}) and PCo1 (estimate $= -0.56$, $P = 0.008$, $r^2 = 0.09$). This result indicates a negative gradient of genetic structuring from spiders that live on the ground level and are active hunters to spiders that build
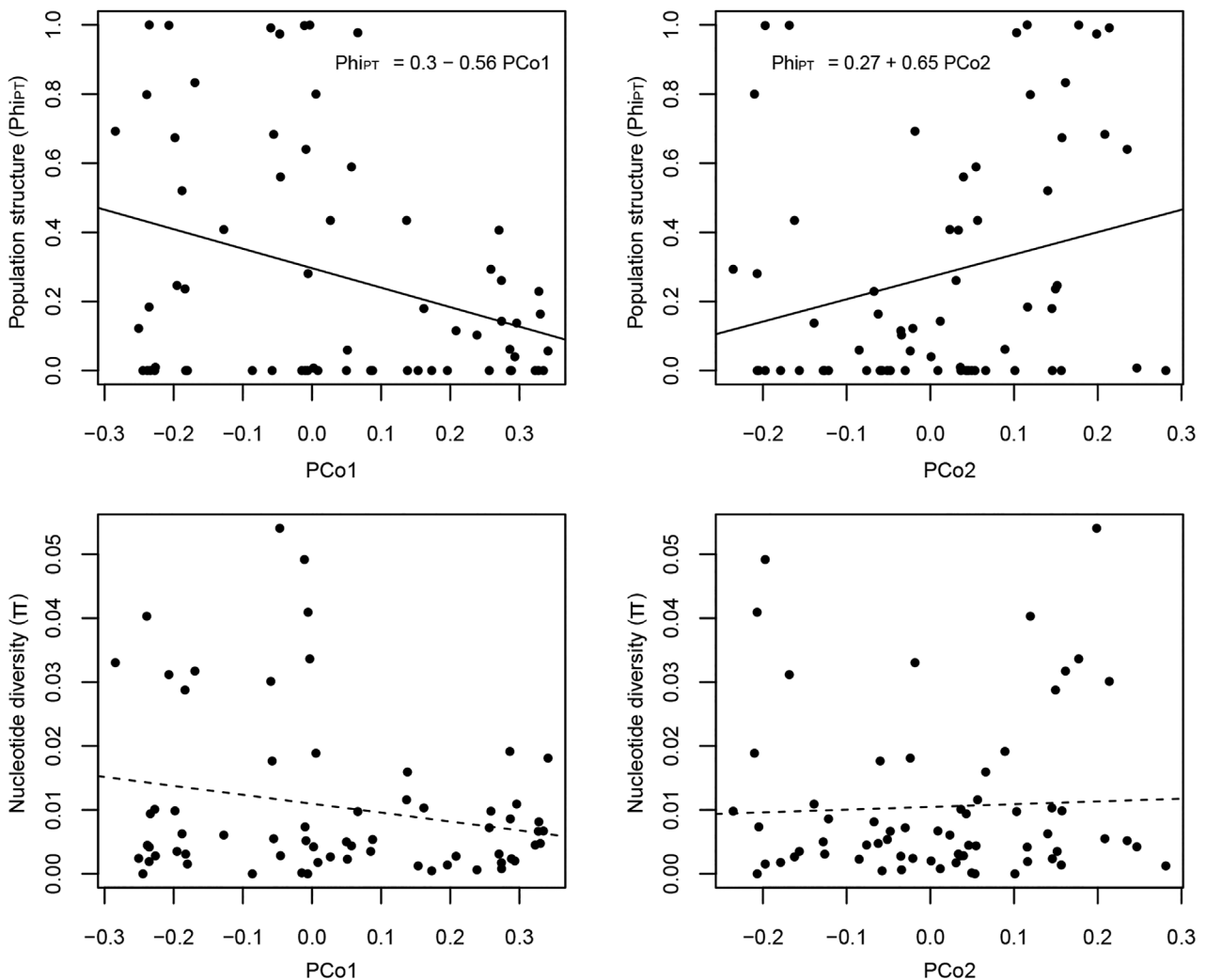

Fig 6. PGLS models correlating population structure (Phi_{PT}) and nucleotide diversity (π) with principal coordinates PCo1 and PCo2. Non-dashed lines and model equations are shown for models with $P < 0.05$.

space webs and live in higher layers of vegetation. Similarly, we also found a positive correlation between genetic population structure and PCo2 (estimate = 0.65, $P = 0.04$, $r^2 = 0.05$). In this case, the positive gradient goes from diurnal to nocturnal species, which are poor dispersers and show high genetic structuring (Fig. 6). Nucleotide diversity (π) was not significantly correlated with either PCo1 ($P = 0.08$) or PCo2 ($P = 0.72$).

Discussion

A barcode library for Iberian spiders

Our study is the largest DNA barcoding study of spiders performed in the Iberian Peninsula to date (3207 barcodes generated) and contributed DNA barcodes of 133 species not previously available. The 371 species for which we provide DNA barcodes represent a quarter of the total number of species recorded in the Iberian Peninsula (Branco *et al.*, 2019). This is especially remarkable given that the Iberian spider diversity is exceptionally large, Spain ranking third among the most diverse countries in Europe, after France and Italy (Nentwig *et al.*, 2017). In addition, Iberian spiders (and communities) are highly endemic (Branco *et al.*, 2019; Malumbres-Olarte *et al.*, 2020). Almost all the analysed species (98.9%) possessed haplotypes that were unique to them, which highlights the need of a local DNA barcode reference library for the effective identification of Iberian species. Moreover, the majority of morphospecies (94%) presented a barcode gap, which is critical for species identification and delimitation by DNA barcoding. In the few cases where our DNA barcode library was unable to provide a species-level identification, it still narrowed it down to two closely related species. These results support the ability of a DNA barcode library to provide accurate identifications for Iberian spiders.

The COI phylogenetic tree (Fig. S2) was well resolved and supported at the species level, with only 16 out of the 371 morphospecies recovered non-monophyletic. As expected from a fast-evolving single mitochondrial marker (Bidegaray-Batista & Arnedo, 2011), deeper nodes were poorly resolved and less supported.

Comparative performance of alternative single markers approaches to species delimitation

Most of the morphospecies matched a single molecular cluster, approximately three quarters in BIN and two-thirds in mPTP. However, while the BIN approach had a 'splitter' tendency (358 morphospecies in 441 BINs), the mPTP model tended to increase the 'merge' cases, generating fewer clusters than predefined morphospecies (371 morphospecies in 365 mPTP clusters). Single-locus species delimitation methods are known to overestimate species diversity because they likely reflect population structure rather than actual species (Carstens *et al.*, 2013; Hawlitschek *et al.*, 2018; Miralles and Vences, 2013). However, the BIN approach has been found to perform better in terms of correspondence with morphologically delimited species

(Blagoev *et al.*, 2013; Huang *et al.*, 2019). If one assumes that morphological variation accurately reflects genome-wide divergence (Muster and Michalik, 2020), then the BINs approach is the single-gene delimitation method that provides a more accurate estimate of the actual species diversity. Also, the proportion of splits in the BIN approach (19.6%) was very similar to the values reported in a similar study conducted on Canadian spiders (19.3%) (Blagoev *et al.*, 2015).

Conversely, the cases of BINs with merged species were infrequent and biologically feasible, inasmuch all the species pairs merged were extremely similar and only recognisable by subtle genitalic differences (Crespo *et al.*, 2018). For example, females of the species *Agyneta rurestris* and *A. pseudorurestris* are morphologically indistinguishable, and males only differ by a little tooth on the lamella of the copulatory bulb. In fact, the specific status of these species has been questioned (Bosmans *et al.*, 2010). In contrast, the mPTP model merged morphospecies that were clearly diagnosable. Several studies conducted in organisms such as golden orb-web spiders (Čandek *et al.*, 2020), trapdoor spiders (Huey *et al.*, 2019) or frogs (Correa *et al.*, 2017) have reported a tendency of the mPTP model to overmerge. Surprisingly, in a few cases, the mPTP model clustered together morphospecies from different genera or even families for unknown reasons. The PTP methods have been shown to be sensitive to factors such as the accuracy of the input tree, population sizes, divergence times or the ratio of population size to divergence time, and to ongoing gene flow (Luo *et al.*, 2018). In our study, the cases of merged morphospecies could be partially explained by the assumption of reciprocal monophyly inherent to the tree-based methods of species delimitation (Fujisawa & Barraclough, 2013), such as mPTP. For instance, the single case of mixture in the mPTP model involved *Neriene clathrata*, whose individuals were nested within those of *Microneta viaria*, rendering the last species paraphyletic. The mPTP split *M. viaria* in six MOTUs, one of them also containing *N. clathrata* sequences.

Finally, the only case of 'mixture' using the BIN method involved three species of cobweb spiders that have been recently shown to have been involved in old introgression events (Domènech *et al.*, 2020).

Uncovering geographic patterns of intraspecific genetic variation across species

A total of 27 morphospecies were consistently split by both BIN and mPTP methods. Interestingly, none of those species were frequent ballooners (see Carvalho & Cardoso, 2014), about half were rare ballooners, and the rest were classified as occasional ballooners. Ballooning is the main means of long-range dispersal in spiders, so it is not surprising that most of the species presenting distinct intraspecific population structure showed a low tendency to balloon.

A negative correlation between morphological disparity and genetic differentiation could hint at the potential existence of cryptic species (Struck *et al.*, 2018). Although single molecular markers may not accurately reflect genome-wide patterns of genetic differentiation, they may identify situations that deserve

further scrutiny. Here, we focused on the cases of splits of single morphospecies into two or more genetic clusters, using IBD plots to reveal the underlying mechanisms of the geographic patterns of genetic distribution. Distinguishing those mechanisms is critical because genetic differentiation in overlapping (sympatric) taxa may indicate reproductive isolation, while genetic divergence between geographically or ecologically distant populations may be better interpreted as population structure rather than species differences.

We identified 13 sympatric and 11 allopatric splits, while in three cases, there were no clear geographic patterns of intraspecific genetic diversity distribution. Among the sympatric splits, we found examples of potential overlooked species. For instance, we observed a single female of the orb weaver *N. umbratica* co-occurring with other conspecifics yet showing high genetic divergence (Fig. 4c). This specimen turned out to be the smallest individual (3.6 mm of prosoma length, the rest ranging from 3.9 to 5.8) and showed slight differences in the epigyne. Other examples of potential overlooked species were found among splits that showed allopatric patterns. For example, we observed that populations of *Eratigena montigena* from the northern parks (Picos de Europa) split into a different cluster from those in the southern parks (Monfragüe and Cabañeros). Morphological re-examination revealed variation in the male palp tibial apophysis and the shape and thickness of the distal end of the conductor that matched the north–south split. Further sampling in intermediate regions and additional sequencing of nuclear genes would be required to confirm the existence of two recently diverged species.

Allopatric differentiation was also found in Sierra Nevada, which suggests a potential role of this mountain system as a phylogeographic break. The guardstone spider *Phrurolinillus tibialis* and the ant spider *Selamia reticulata* split into different clusters corresponding to the southern (PS1) and northern (PS2) slopes of Sierra Nevada (in the case of *Selamia*, the last cluster also included individuals from the central parks). Although the two plots were barely 30 km apart, they were separated by elevations above 2500 m, which may constitute an insurmountable barrier for the dispersal of these taxa.

Population structure is associated with functional traits

Although genetic diversity (i.e. nucleotide diversity of mtDNA) and functional traits were decoupled, we uncovered a significant correlation between those traits and population structure. In general, active hunters, ground-dwellers and nocturnal species tend to be more genetically structured, while species living in higher layers of the vegetation, those mostly using space webs to capture prey, those active at daylight and good ballooners, were less genetically structured. Our analyses revealed two groups of highly correlated functional traits. The first group of traits included the active hunter and space web foraging strategies and preferred microhabitats 'ground level', 'herbaceous' and 'very low canopy'. The fact that the active hunting strategy co-varied with the ground dwelling habitat probably reflects a higher efficiency of this foraging strategy in covering larger areas in the ground. Conversely, space webs seemed to co-vary

with 'herbaceous' and 'very low canopy' microhabitats found at the medium vegetation level. Although some webs such as those of some theridiids are located almost at ground level and have threads that attach to the ground and that aim at wandering prey, most spatial webs are built at a certain distance from the ground and are intended to capture flying insects (Cardoso *et al.*, 2011). The second group of traits included the circadian activity and the dispersal ability by ballooning. A higher tendency to balloon was associated with diurnal spiders, possibly because high air temperatures and low wind velocities are particularly important for the initiation of ballooning dispersal (Vugts & Van Wingerden, 1976; Reynolds *et al.*, 2007). Although whether wind speed is higher or lower during daytime is not clear and might depend on the season (Pérez *et al.*, 2004; Emeis *et al.*, 2007), temperature is usually warmer during daylight, which may explain the association between ballooning and diurnality. We did not recover a significant relationship between ballooning tendency and small body size. This is hardly surprising inasmuch most studies supporting this relationship were based on the measurements of the individuals dispersing by ballooning, in many cases including immatures (Dean & Sterling, 1985; Greenstone *et al.*, 1987), while we used the measurements of the adult stage of each species. Ballooning dispersal is not restricted to spider species with a small adult size because the juvenile stages of spiders with large adult sizes can also balloon (Humphrey, 1987; Bonte *et al.*, 2003b).

The two groups of highly correlated functional traits showed significant phylogenetic signal. These results are congruent with the observation that most of these traits are known to be highly conserved at the genus or even family levels (Cardoso *et al.*, 2011).

The deep population structure recovered in active ground dwelling hunters could be explained by the high habitat fidelity of ground species. Such habitat fidelity would result in more fragmented suitable habitat patches, and consequently less connected populations, as has been observed in the purse-web spider *Atypus affinis* Eichwald, 1830 (Pétillon *et al.*, 2012). In fact, Malumbres-Olarte *et al.* (2020) found ground-dwelling species to have higher levels of endemicity than species living higher in the vegetation. Interestingly, the ballooning capacity was found to be positively correlated with the PCo1, just like vegetation layers 'herbaceous' and 'very low canopy' and opposite to the tendency to live on the ground. Blandenier (2009) found ground-dwelling spiders to be present in ballooning captures almost in the same proportion as spiders living in the vegetation, although ground spiders were rare in the case of species from forested areas. Bell *et al.* (2005) suggested that species living at ground level may be less exposed to wind and convective currents compared to spiders living higher in the vegetation, but it is also possible that ground-dwelling spiders simply climb to higher vegetation layers when they need to disperse by ballooning. Also, spider populations appear to be more genetically structured in species that are nocturnal and do not use ballooning for dispersal and less genetically structured in species active at daylight. There is little evidence of a direct connection between the circadian activity and population structure. Most likely, whether spiders are diurnal or nocturnal explains population structure indirectly by affecting the ballooning ability. As the main means

of long-range dispersal in spiders, ballooning is known to contribute importantly to the connectivity and gene flow among their populations (Bonte *et al.*, 2003a; Dimassi *et al.*, 2016, but see Ramirez & Haakonsen, 1999), thus affecting the levels of genetic population structure and the level of endemicity (Malumbres-Olarte *et al.*, 2020). Although many spiders are able to disperse by walking, mostly in search of mating partners, their impact to connect populations at the geographical scale of this study (from 0.11 to 721 km) is probably negligible. The positive relationship between ground dwellers and population structure was further supported by the fact that species involved in 'splits' mostly included poor or no ballooners, so the lower levels of airborne dispersal would contribute to reducing the gene flow between populations.

Our former discussion was based on the assumption that mtDNA information accurately reflects population history and demography, which may not always be true. Mitochondrial genes, such as COI, are maternally inherited, so they may only reflect the population structure of females in species with different dispersal behaviours between sexes. Additionally, ecological and evolutionary processes such as incomplete lineage sorting or hybridization may negatively affect the results of single-gene population structure analyses, which could be minimised by using additional markers. Also, the modest number of localities and of individuals sampled per species could also affect the estimates of genetic diversity and population structure and ultimately the mechanisms described therein.

Conclusions

Here, we present the most extensive DNA barcoding study conducted on Iberian spiders to date. The assembled DNA barcode library will not only ease the identification of single specimens, but it will also facilitate the automatisation of the identification of bulk samples using high-throughput sequencing (HTS) approaches (e.g. metabarcoding, genomic skimming) for large-scale cataloguing projects. Because of the emergency for the rapid inventory and monitoring of the declining biodiversity, the use of HTS approaches is becoming more widespread. However, morphologically validated reference sequences are still essential to secure identification and to link the wealth of genetic data to name-based heritage information. Here, the distance-based delimitation method (BIN) outperformed the character-based method (mPTP) in terms of matching to morphospecies reference. The patterns of geographic mtDNA variation revealed by IBD plots allowed us to uncover instances of overlooked diversity but also of potential phylogeographic breaks. Our multi-taxon comparison of mtDNA variation suggested that functional traits of spiders, specifically active hunting and ground dwelling, as well as nocturnal activity and ballooning ability may partially explain population structure. The large-scale patterns of genetic variation in spider communities across the Spanish National Park network revealed in our study and their relationship to species traits will contribute important information to evaluate and guide future conservation and management strategies.

Conflict of Interest

All the authors declare that they have no conflict of interest. There are no disputes over the ownership of the data presented in the paper and all contributions have been attributed appropriately.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon request.

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Fig. S1 Barcode gap for the 371 species of Iberian spiders with two or more sampled individuals. Maximum intraspecific divergence is plotted against nearest-neighbour distance. Grey triangles represent morphospecies consistently merged in both BIN and mPTP, while black circles show the rest of possible combinations. Points above the diagonal line indicate species with a barcode gap.

Fig. S2 Maximum likelihood tree of all the sequences of the COI mitochondrial gene. Numbers in the nodes indicate the values of ML ultrafast Bootstrap support.

Fig. S3 Isolation by Distance (IBD) plots of the genetic distance (%) in relation to the geographic distance (km) of all morphospecies split in both delimitation methods. Intra- and intercluster distances according to BIN delimitation are represented with triangles and circles, respectively. Intra- and intercluster distances according to mPTP delimitation are represented in blue and orange, respectively.

Table S1 Categorization of the traits used in this study and their corresponding states.

Table S2 Values of Spearman correlation between each trait and PCo1 and PCo2.

Table S3 List of the 67 species used for population structure and genetic diversity analyses with the number of individuals

included from each site (and the total). Species are sorted by total abundance. PP: Picos de Europa; PO: Ordesa y Monte Perdido; PA: Aigüestortes i Estany de Sant Maurici; PM: Monfragüe; PC: Cabañeros; PS: Sierra Nevada.

REFERENCES

Astrin, J.J., Höfer, H., Spelda, J., Holstein, J., Bayer, S., Hendrich, L., Huber, B.A., Kielhorn, K., Krammer, H., Lemke, M., Monje, J.C., Mornière, J., Rulik, B., Petersen, M., Janssen, H. & Muster, C. (2016) Towards a DNA Barcode Reference Database for Spiders and Harvestmen of Germany. *PLoS ONE*, **11**(9), 1–24. https://doi.org/10.1371/journal.pone.0162624.

Barlow, J., Lennox, G., Ferreira, J., Berenguer, E., Lees, A., Mac Nally, R., Thomson, J., Ferraz, S., Louzada, J., Oliveira, V., Parry, L., Solar, R., Guimarães Vieira, I., Aragão, L., Begotti, R., Braga, R., Cardoso, T., Oliveira Junior, R., Souza, C. & Gardner, T. (2016) Anthropogenic disturbance in tropical forests can double biodiversity loss from deforestation. *Nature*, **535**(7610), 144–147. https://doi.org/10.1038/nature18326.

Barrett, R.D.H. & Hebert, P. (2005) Identifying spiders through DNA barcodes. *Canadian Journal of Zoology*, **83**(3), 481–491. https://doi.org/10.1139/z05-024.

Bell, J.R., Bohan, D.A., Shaw, E.M. & Weyman, G.S. (2005) Ballooning dispersal using silk: world fauna, phylogenies, genetics and models. *Bulletin of Entomological Research*, **95**, 69–114.

Benson, D.A., Cavanaugh, M., Clark, K., Karsch-Mizrachi, I., Lipman, D.J., Ostell, J. & Sayers, E.W. (2017) GenBank. *Nucleic Acids Research*, **45**(D1), D37–D42. https://doi.org/10.1093/nar/gkw1070.

Bidegaray-Batista, L. & Arnedo, M.A. (2011) Gone with the plate: the opening of the Western Mediterranean basin drove the diversification of ground-dweller spiders. *BMC Evolutionary Biology*, **11**(1), 317. https://doi.org/10.1186/1471-2148-11-317.

Blagoev, G.A., de Waard, J.R., Ratnasingham, S., de Waard, S.L., Lu, L., Robertson, J., Telfer, A.C. & Hebert, P.D.N. (2015) Untangling taxonomy: a DNA barcode reference library for Canadian spiders. *Molecular Ecology Resources*, **16**(1), 325–341. https://doi.org/10.1111/1755-0998.12444.

Blagoev, G.A., Nikolova, N.I., Sobel, C.N., Hebert, P. & Adamowicz, S.J. (2013) Spiders (Araneae) of Churchill, Manitoba: DNA barcodes and morphology reveal high species diversity and new Canadian records. *BMC Ecology*, **13**, 1–17.

Blandenier, G. (2009) Ballooning of spiders (Araneae) in Switzerland: general results from an eleven-year survey. *Arachnology*, **14**(7), 308–316. https://doi.org/10.13156/arac.2009.14.7.308.

Bonte, D., Baert, L., Lens, L. & Maelfait, J.P. (2004) Effects of aerial dispersal, habitat specialisation, and landscape structure on spider distribution across fragmented grey dunes. *Ecography*, **27**(3), 343–349. https://doi.org/10.1111/j.0906-7590.2004.03844.x.

Bonte, D., Lens, L., Maelfait, J.-P., Hoffmann, M. & Kuijken, E. (2003a) Patch quality and connectivity influence spatial dynamics in a dune wolfspider. *Oecologia*, **227–233**, 227–233. https://doi.org/10.1007/s00442-003-1195-6.

Bonte, D., Vandenbroecke, N., Lens, L. & Maelfait, J.-P. (2003b) Low propensity for aerial dispersal in specialist spiders from fragmented landscapes. *Proceedings of the Royal Society B: Biological Sciences*, **270**(1524), 1601–1607. https://doi.org/10.1098/rspb.2003.2432.

Bosmans, R., Cardoso, P. & Crespo, L.C. (2010) A review of the linyphiid spiders of Portugal, with the description of six new species (Araneae: Linyphiidae). *Zootaxa*, **2473**, 1.

Botham, J.L., Haddad, C.R., Gryzenhout, M., Swart, V.R. & Bredenhand, E. (2020) High genetic diversity of spider species in a mosaic montane grassland landscape. *PLoS One*, **15**(6), 1–26. https://doi.org/10.1371/journal.pone.0234437.

Branco, V.V., Morano, E. & Cardoso, P. (2019) An update to the Iberian spider checklist (Araneae). *Zootaxa*, **4614**(2), 201–254. https://doi.org/10.11646/zootaxa.4614.2.1.

Čandek, K., Agnarsson, I., Binford, G.J. & Kuntner, M. (2020) Caribbean golden orbweaving spiders maintain gene flow with North America. *Zoologica Scripta*, **49**(2), 210–221. https://doi.org/10.1111/zsc.12405.

Cardoso, P. (2009) Standardization and optimization of arthropod inventories-the case of Iberian spiders. *Biodiversity and Conservation*, **18**(14), 3949–3962. https://doi.org/10.1007/s10531-009-9690-7.

Cardoso, P., Pekár, S., Jocqué, R. & Coddington, J.A. (2011) Global patterns of guild composition and functional diversity of spiders. *PLoS ONE*, **6**(6), e21710. https://doi.org/10.1371/journal.pone.0021710.

Carstens, B.C., Pelletier, T.A., Reid, N.M. & Satler, J.D. (2013) How to fail at species delimitation. *Molecular Ecology*, **22**(17), 4369–4383. https://doi.org/10.1111/mec.12413.

Carvalho, J.C. & Cardoso, P. (2014) Drivers of beta diversity in Macaronesian spiders in relation to dispersal ability. *Journal of Biogeography*, **41**(10), 1859–1870. https://doi.org/10.1111/jbi.12348.

Casiraghi, M., Labra, M., Ferri, E., Galimberti, A. & de Mattia, F. (2010) DNA barcoding: a six-question tour to improve users' awareness about the method. *Briefings in Bioinformatics*, **11**(4), 440–453. https://doi.org/10.1093/bib/bbq003.

Chapman, A.D. (2009) *Numbers of Living Species in Australia and the World*, Vol. **2**, p. 84. Toowoomba, Australia: Heritage. https://doi.org/10.1177/135.

Charlesworth, D. & Willis, J.H. (2009) The genetics of inbreeding depression. *Nature Reviews Genetics*, **10**(11), 783–796. https://doi.org/10.1038/nrg2664.

Chernomor, O., von Haeseler, A. & Minh, B.Q. (2016) Terrace aware data structure for phylogenomic inference from supermatrices. *Systematic Biology*, **65**, 997–1008. https://doi.org/10.1093/sysbio/syw037.

Correa, C., Vásquez, D., Castro-Carrasco, C., Zúñiga-Reinoso, Á., Ortiz, J.C. & Palma, R.E. (2017) Species delimitation in frogs from South American temperate forests: the case of *Eupsophus*, a taxonomically complex genus with high phenotypic variation. *PLoS ONE*, **12**(8), 1–21. https://doi.org/10.1371/journal.pone.0181026.

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., Opatova, 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*, **6**, e29443 1–273. https://doi.org/10.3897/BDJ.6.e29443.

Dean, D.A. & Sterling, W.L. (1985) Size and phenology of ballooning spiders at two locations in Eastern Texas. *American Arachnological Society*, **13**(1), 111–120.

Dimassi, N., Khadra, Y., Othmen, A.B., Ezzine, I.K. & Said, K. (2016) High genetic diversity vs. low genetic and morphological differentiation of *Argiope trifasciata* (Araneae, Araneidae) in Tunisia. *Systematics and Biodiversity*, **15**, 1–15. https://doi.org/10.1080/14772000.2016.1203040.

Doherty, T.S., Glen, A.S., Nimmo, D.G., Ritchie, E.G. & Dickman, C.R. (2016) Invasive predators and global biodiversity loss. *Proceedings of the National Academy of Sciences of the United States of America*, **113**(40), 11261–11265. https://doi.org/10.1073/pnas.1602480113.

Domènech, M., Crespo, L.C., Enguídanos, A. & Arnedo, M.A. (2020) Mitochondrial discordance in closely related *Theridion* spiders (Araneae, Theridiidae), with description of a new species of the *T. melanurum* group. *Zoosystematics and Evolution*, **96**(1), 159–173. https://doi.org/10.3897/zse.9.

Dray, S. & Dufour, A.B. (2007) The ade4 package: implementing the duality diagram for ecologists. *Journal of Statistical Software*, **22**(4), 1–20. https://doi.org/10.18637/jss.v022.i04.

Elias, M., Joron, M., Willmott, K., Silva-Brandão, K.L., Kaiser, V., Arias, C.F., Piñerez, L.M.G., Uribe, S., Brower, A.V.Z., Freitas, A.V.L. & Jiggins, C.D. (2009) Out of the Andes: patterns of diversification in clearwing butterflies. *Molecular Ecology*, **18**(8), 1716–1729. https://doi.org/10.1111/j.1365-294X.2009.04149.x.

Ellegren, H. & Galtier, N. (2016) Determinants of genetic diversity. *Nature Reviews Genetics*, **17**(7), 422–433. https://doi.org/10.1038/nrg.2016.58.

Emeis, S., Baumann-Stanzer, K., Piringer, M., Kallistratova, M., Kouznetsov, R. & Yuhkov, V. (2007) Wind and turbulence in the urban boundary layer - Analysis from acoustic remote sensing data and fit to analytical relations. *Meteorologische Zeitschrift*, **16**(4), 393–406. https://doi.org/10.1127/0941-2948/2007/0217.

Fernández, R., Kallal, R.J., Dimitrov, D., Ballesteros, J.A., Arnedo, M.A., Giribet, G. & Hormiga, G. (2018) Phylogenomics, diversification dynamics, and comparative transcriptomics across the spider tree of life. *Current Biology*, **28**(9), 1489–1497.e5. https://doi.org/10.1016/j.cub.2018.03.064.

Folmer, O., Black, M., Hoeh, W., Lutz, R. & Vrijenhoek, R. (1994) DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology*, **3**, 294–299.

Fujisawa, T. & Barraclough, T.G. (2013) Delimiting species using single-locus data and the generalized mixed yule coalescent approach: a revised method and evaluation on simulated data sets. *Systematic Biology*, **62**(5), 707–724. https://doi.org/10.1093/sysbio/syt033.

Greenstone, M.H., Morgan, C.E. & Hultsch, A.-L. (1987) Ballooning spiders in Missouri, USA, and New South Wales, Australia: family and mass distributions. *Journal of Arachnology*, **15**, 163–170. https://doi.org/10.2307/3705725.

Hajibabaei, M., Shokralla, S., Zhou, X., Singer, G.A.C. & Baird, D.J. (2011) Environmental barcoding: a next-generation sequencing approach for biomonitoring applications using river benthos. *PLoS One*, **6**(4), 1–7. https://doi.org/10.1371/journal.pone.0017497.

Hamilton, C.A., Hendrixson, B.E., Brewer, M.S. & Bond, J.E. (2014) An evaluation of sampling effects on multiple DNA barcoding methods leads to an integrative approach for delimiting species: a case study of the North American tarantula genus *Aphonopelma* (Araneae, Mygalomorphae, Theraphosidae). *Molecular Phylogenetics and Evolution*, **71**(1), 79–93. https://doi.org/10.1016/j.ympev.2013.11.007.

Hawlitschek, O., Scherz, M.D., Ruthensteiner, B., Crottini, A. & Glaw, F. (2018) Computational molecular species delimitation and taxonomic revision of the gecko genus *Ebenavia* Boettger, 1878. *Science of Nature*, **105**(7–8), 49. https://doi.org/10.1007/s00114-018-1574-9.

Huang, X.C., Su, J.H., Ouyang, J.X., Ouyang, S., Zhou, C.H. & Wu, X.P. (2019) Towards a global phylogeny of freshwater mussels (Bivalvia: Unionida): species delimitation of Chinese taxa, mitochondrial phylogenomics, and diversification patterns. *Molecular Phylogenetics and Evolution*, **130**, 45–59. https://doi.org/10.1016/j.ympev.2018.09.019.

Huey, J.A., Hillyer, M.J. & Harvey, M.S. (2019) Phylogenetic relationships and biogeographic history of the Australian trapdoor spider genus *Conothele* (Araneae: Mygalomorphae: Halonoproctidae): diversification into arid habitats in an otherwise tropical radiation.

Invertebrate Systematics, **33**(4), 628–643. https://doi.org/10.1071/IS18078.

Humphrey, J.A.C. (1987) Fluid mechanic constraints on spider ballooning. *Oecologia*, **73**(3), 469–477. https://doi.org/10.1007/BF00385267.

Jiménez-Valverde, A., Baselga, A., Melic, A. & Txasko, N. (2010) Climate and regional beta-diversity gradients in spiders: dispersal capacity has nothing to say? *Insect Conservation and Diversity*, **3**(1), 51–60. https://doi.org/10.1111/j.1752-4598.2009.00067.x.

Kapli, P., Lutteropp, S., Zhang, J., Kobert, K., Pavlidis, P., Stamatakis, A. & Flouri, T. (2017) Multi-rate Poisson tree processes for single-locus species delimitation under maximum likelihood and Markov chain Monte Carlo. *Bioinformatics*, **33**(11), 1630–1638. https://doi.org/10.1093/bioinformatics/btx025.

Katoh, K., Rozewicki, J. & Yamada, K.D. (2019) MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Briefings in Bioinformatics*, **20**(4), 1160–1166.

Kennedy, S.R., 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.

Kumar, S., Stecher, G. & Tamura, K. (2016) MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Molecular Biology and Evolution*, **33**(7), 1870–1874. https://doi.org/10.1093/molbev/msw054.

Luo, A., Ling, C., Ho, S.Y.W. & Zhu, C.D. (2018) Comparison of methods for molecular species delimitation across a range of speciation scenarios. *Systematic Biology*, **67**(5), 830–846. https://doi.org/10.1093/sysbio/syy011.

Macías-Hernández, N., Domènech, M., Cardoso, P., Emerson, B.C., Borges, P.A.V., Lozano-Fernandez, J., Paulo, O.S., Vieira, A., Enguídanos, A., Rigal, F., Amorim, I.R. & Arnedo, M.A. (2020a) Building a robust, densely-sampled spider tree of life for ecosystem research. *Diversity*, **12**(288), 1–23.

Macías-Hernández, N., Ramos, C., Domènech, M., Febles, S., Santos, I., Arnedo, M.A., Borges, P.A.V., Emerson, B.C. & Cardoso, P. (2020b) A database of functional traits for spiders from native forests of the Iberian Peninsula and Macaronesia. *Biodiversity Data Journal*, **8**, 1–12. https://doi.org/10.3897/bdj.8.e49159.

Malumbres-Olarte, J., Crespo, L.C., Domènech, M., Cardoso, P., Moya-Laraño, J., Ribera, C. & Arnedo, M.A. (2020) How Iberian are we? Mediterranean climate determines structure and endemicity of spider communities in Iberian oak forests. *Biodiversity and Conservation*, **29**, 3973–3996. https://doi.org/10.1007/s10531-020-02058-7.

Meirmans, P.G. & Hedrick, P.W. (2011) Assessing population structure: FST and related measures. *Molecular Ecology Resources*, **11**(1), 5–18. https://doi.org/10.1111/j.1755-0998.2010.02927.x.

Miller, M.A., Pfeiffer, W. & Schwartz, T. (2010) Creating the CIPRES Science Gateway for inference of large phylogenetic trees. 2010 Gateway Computing Environments Workshop, GCE 2010 https://doi.org/10.1109/GCE.2010.5676129.

Miralles, A. & Vences, M. (2013) New metrics for comparison of taxonomies reveal striking discrepancies among species delimitation methods in *Madascincus* lizards. *PLoS One*, **8**(7), e68242. https://doi.org/10.1371/journal.pone.0068242.

Morley, E.L. & Robert, D. (2018) Electric fields elicit ballooning in spiders. *Current Biology*, **28**(14), 2324–2330.e2. https://doi.org/10.1016/j.cub.2018.05.057.

Múrria, C., Rugenski, A.T., Whiles, M.R. & Vogler, A.P. (2015) Long-term isolation and endemicity of neotropical aquatic insects limit the community responses to recent amphibian decline. *Diversity and Distributions*, **21**(8), 938–949. https://doi.org/10.1111/ddi.12343.

Muster, C. & Michalik, P. (2020) Cryptic diversity in ant-mimic *Micaria* spiders (Araneae, Gnaphosidae) and a tribute to early naturalists. *Zoologica Scripta*, **49**, 197–209. https://doi.org/10.1111/zsc.12404.

Nentwig, W., Blick, T., Bosmans, R., Gloor, D., Hänggi, A., Kropf, C. (2017) Spiders of Europe. Version 5.2018. <https://www.araneae.nmbe.ch/> 21st May 2018

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. doi:10.1093/molbev/msu300.

Nyffeler, M. & Birkhofer, K. (2017) An estimated 400-800 million tons of prey are annually killed by the global spider community. *Science of Nature*, **104**(3–4), 30. https://doi.org/10.1007/s00114-017-1440-1.

Orme, D., Freckleton, R.P., Thomas, G.H., Petzoldt, T., Fritz, S.A. & Bendall Isaac, N.J. (2013) CAPER: comparative analyses of phylogenetics and evolution in R. *Methods in Ecology and Evolution*, **3**, 145–151.

Pagel, M. (1999) Inferring the historical patterns of biological evolution. *Nature*, **401**(6756), 877–884. https://doi.org/10.1038/44766.

Papadopoulou, A., Anastasiou, I., Keskin, B. & Vogler, A.P. (2009) Comparative phylogeography of tenebrionid beetles in the Aegean archipelago: the effect of dispersal ability and habitat preference. *Molecular Ecology*, **18**(11), 2503–2517. https://doi.org/10.1111/j.1365-294X.2009.04207.x.

Paradis, E. (2010) pegas: an R package for population genetics with an integrated–modular approach. *Bioinformatics*, **26**, 419–420.

Paradis, E. & Schliep, K. (2018) ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics*, **35**, 526–528.

Pérez, I.A., García, M.A., Sánchez, M.L. & de Torre, B. (2004) Analysis of height variations of sodar-derived wind speeds in Northern Spain. *Journal of Wind Engineering and Industrial Aerodynamics*, **92**, 875–894. https://doi.org/10.1016/j.jweia.2004.05.002.

Peters, R.H. & Wassenberg, K. (1983) The effect of body size on animal abundance. *Oecologia*, **60**, 89–96. https://doi.org/10.1007/BF00379325.

Pétillon, J., Deruytter, D., Decae, A., Renault, D. & Bonte, D. (2012) Habitat use, but not dispersal limitation, as the mechanism behind the aggregated population structure of the mygalomorph species *Atypus affinis*. *Animal Biology*, **62**(2), 181–192. https://doi.org/10.1163/157075611X617094.

Pimm, S.L., Jenkins, C.N., Abell, R., Brooks, T.M., Gittleman, J.L., Joppa, L.N., Raven, P.H. & Sexton, J.O. (2014) The biodiversity of species and their rates of extinction, distribution, and protection. *Science*, **344**(6187), 987–1005. https://doi.org/10.1126/science.1246752.

Pope, L.C., Domingo-Roura, X., Erven, K. & Burke, T. (2006) Isolation by distance and gene flow in the Eurasian badger (*Meles meles*) at both a local and broad scale. *Molecular Ecology*, **15**(2), 371–386. https://doi.org/10.1111/j.1365-294X.2005.02815.x.

R Core Team (2020) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria https://www.R-project.org/.

Ramírez, M.G. & Chi, B. (2004) Cryptic speciation, genetic diversity and gene flow in the California turret spider Atypoides riversi (Araneae: Antrodiaetidae). *Biological Journal of the Linnean Society*, **82**, 27–37.

Ramirez, M.G. & Haakonsen, K.E. (1999) Gene flow among habitat patches on a fragmented landscape in the spider *Argiope trifasciata* (Araneae: Araneidae). *Heredity*, **83**(5), 580–585. doi:10.1038/sj.hdy.6885630.

Ramsay, S.L. & Houston, D.C. (2003) Amino acid composition of some woodland arthropods and its implications for breeding tits and other passerines. *Ibis*, **145**, 227–232.

Ratnasingham, S. & Hebert, P. (2013) A DNA-Based Registry for All Animal Species: The Barcode Index Number (BIN) System.

PLoS ONE, **8**(8), 1–16. https://doi.org/10.1371/journal.pone.0066213.

Revell, L.J. (2012) phytools: an R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution*, **3**(2), 217–223. https://doi.org/10.1111/j.2041-210X.2011.00169.x.

Reynolds, A.M., Bohan, D.A. & Bell, J.R. (2007) Ballooning dispersal in arthropod taxa: conditions at take-off. *Biology Letters*, **3**(3), 237–240. doi:10.1098/rsbl.2007.0109.

Salinas-Ivanenko, S. & Múrria, C. (2021) Macroecological trend of increasing values of intraspecific genetic diversity and population structure from temperate to tropical streams. *Global Ecology and Biogeography*, **30**, 1685–1697. doi:10.1111/geb.13344.

Schäfer, M.A., Hille, A. & Uhl, G.B. (2001) Geographical patterns of genetic subdivision in the cellar spider *Pholcus phalangioides* (Araneae). *Heredity*, **86**(1), 94–102. https://doi.org/10.1046/j.1365-2540.2001.00815.x.

Settepani, V., Bechsgaard, J. & Bilde, T. (2014) Low genetic diversity and strong but shallow population differentiation suggests genetic homogenization by metapopulation dynamics in a social spider. *Journal of Evolutionary Biology*, **27**, 2850–2855.

Slatkin, M. (1987) Gene flow and the geographic structure of natural populations. *Science*, **236**(4803), 787–792. https://doi.org/10.1126/science.3576198.

Socolar, J.B., Valderrama Sandoval, E.H. & Wilcove, D.S. (2019) Overlooked biodiversity loss in tropical smallholder agriculture. *Conservation Biology*, **33**(6), 1338–1349. https://doi.org/10.1111/cobi.13344.

Spooner, D.M. (2009) DNA barcoding will frequently fail in complicated groups: an example in wild potatoes. *American Journal of Botany*, **96**(6), 1177–1189. https://doi.org/10.3732/ajb.0800246.

Struck, T.H., Feder, J.L., Bendiksby, M., Birkeland, S., Cerca, J., Gusarov, V.I., Kistenich, S., Larsson, K.H., Liow, L.H., Nowak, M.D., Stedje, B., Bachmann, L. & Dimitrov, D. (2018) Finding evolutionary processes hidden in cryptic species. *Trends in Ecology and Evolution*, **33**(3), 153–163. https://doi.org/10.1016/j.tree.2017.11.007.

Tyagi, K., Kumar, V., Kundu, S., Pakrashi, A., Prasad, P., Caleb, J.T.D. & Chandra, K. (2019) Identification of Indian Spiders through DNA barcoding: cryptic species and species complex. *Scientific Reports*, **9**(1), 1–13. https://doi.org/10.1038/s41598-019-50510-8.

Victor, B.C., Hanner, R., Shivji, M., Hyde, J. & Caldow, C. (2009) Identification of the larval and juvenile stages of the Cubera Snapper, *Lutjanus cyanopterus*, using DNA barcoding. *Zootaxa*, **36**, 24–36.

Vugts, H.F. & Van Wingerden, W.K.R.E. (1976) Meteorological aspects of aeronautic behaviour of spiders. *Oikos*, **27**, 433–444.

World Spider Catalog (2020) World Spider Catalog, version 21.5. Natural History Museum Bern, <http://wsc.nmbe.ch/> December 2020

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**(4), 613–623. https://doi.org/10.1111/j.2041-210X.2012.00198.x.

Zeale, M.R.K., Butlin, R.K., Barker, G.L.A., Lees, D.C. & Jones, G. (2011) Taxon-specific PCR for DNA barcoding arthropod prey in bat faeces. *Molecular Ecology Resources*, **11**(2), 236–244. https://doi.org/10.1111/j.1755-0998.2010.02920.x.

Zhang, J., Kapli, P., Pavlidis, P. & Stamatakis, A. (2013) A general species delimitation method with applications to phylogenetic placements. *Bioinformatics*, **29**(22), 2869–2876. https://doi.org/10.1093/bioinformatics/btt499.