Towards a Canary Islands barcode database for soil biodiversity: revealing cryptic and unrecorded mite species diversity within insular soils

Irene Santos-Perdomo[‡,§], Daniel Suárez[‡,§], María L. Moraza[|], Paula Arribas[‡], Carmelo Andújar[‡]

‡ Island Ecology and Evolution Research Group, Institute of Natural Products and Agrobiology (IPNA-CSIC), 38206, La Laguna, Spain
§ School of Doctoral and Postgraduate Studies, University of La Laguna, 38206, La Laguna, Spain
| Universidad de Navarra, Instituto de Biodiversidad y Medioambiente BIOMA, Irunlarrea 1, 31008, Pamplona, Spain

Abstract

Soil arthropod diversity contributes to a high proportion of the total biodiversity on Earth. However, most soil arthropods are still undescribed, hindering our understanding of soil functioning and global biodiversity estimations. Inventorying soil arthropods using conventional taxonomical approaches is particularly difficult and costly due to the great species richness, abundance and local-scale heterogeneity of mesofauna communities and the poor taxonomic background knowledge of most lineages. To alleviate this situation, we have designed and implemented a molecular barcoding framework adapted to soil fauna. This pipeline includes different steps, starting with a morphology-based selection of specimens which are imaged. Then, DNA is extracted non-destructively. Both images and voucher specimens are used to assign a taxonomic identification, based on morphology that is further checked for consistency with molecular information. Using this procedure, we studied 239 specimens of mites from the Canary Islands including representatives of Mesostigmata, Sarcoptiformes and Trombidiformes, of which we recovered barcode sequences for 168 specimens that were morphologically identified to 49 species, with nine specimens that could only be identified at the genus or family levels. Multiple species

delimitation analyses were run to compare molecular delimitations with morphological identifications, including ASAP, mlPTP, BINs and 3% and 8% genetic distance thresholds. Additionally, a species-level search was carried out at the Biodiversity Databank of the Canary Islands (BIOTA) to evaluate the number of species in our dataset that were not previously recorded in the archipelago. In parallel, a sequence-level search of our sequences was performed against BOLD Systems. Our results reveal that multiple morphologically identified species correspond to different molecular lineages, which points to significant levels of unknown cryptic diversity within the archipelago. In addition, we evidenced that multiple species in our dataset constituted new records for the Canary Islands fauna and that the information for these lineages within online genetic repositories is very incomplete. Our study represents the first systematic effort to catalogue the soil arthropod mesofauna of the Canary Islands and establishes the basis for the Canary Islands Soil Biodiversity barcode database. This resource will constitute a step forward in the knowledge of these arthropods in a region of special interest.

Keywords

Acari, COI, barcoding, soil mesofauna, biotic frontier, species inventory, oceanic islands, species delimitation methods

Introduction

Soils harbour a vast proportion of total biodiversity on Earth (Decaëns et al. 2006), which play a crucial role in critical processes, such as soil formation, nutrient and water cycling, climate regulation, production of food, medicine and fibre, disease and pest control (Bardgett and Van Der Putten 2014). At the same time, the soil environment is considered one of the last biotic frontiers to human knowledge (André et al. 1994). Current knowledge gaps regarding soil biodiversity are still massive (Guerra et al. 2020), primarily because of the methodological and logistical complexity of approaching complex soil communities. A striking and concerning consequence of these knowledge gaps is the difficulty in developing and implementing conservation strategies to preserve soil biodiversity (Veresoglou et al. 2015), a direct impediment to the 2030 UN Agenda for Sustainable Development. Recently, the Convention on Biological Diversity and the FAO-Global Soil Partnership actively promoted diverse initiatives to fill critical knowledge gaps for soil biodiversity, starting with the first report on the State of Knowledge of Soil Biodiversity (FAO et al. 2020).

The methodological and logistical issues that have hindered our understanding of soil biodiversity are particularly exacerbated for some edaphic groups and geographical areas (Guerra et al. 2020). A substantial fraction of soil biodiversity is represented by the soil mesofauna, i.e. small-bodied invertebrates measuring between 0.1 and 2 mm that are regularly found in their thousands in every square metre of soil (Walter and Proctor 2013, Nielsen et al. 2015). Soil arthropod mesofauna is a functionally important component of soil communities, directly affecting the physicochemical and biological properties of the soil and

leaf litter (Ponge 2013). However, identifying and quantifying soil arthropod mesofauna using conventional taxonomical approaches is difficult and costly. This is due to the great species richness, abundance and local-scale heterogeneity of mesofauna communities, the minute size of specimens, the scarcity of taxonomic experts and the poor taxonomic background knowledge of most soil lineages (Bardgett 2002, Decaëns 2010). Besides, soil arthropod diversity estimations using morphological techniques could be substantially biased, at least in some lineages, due to the high prevalence of cryptic diversity within morphologically defined species (e.g. Cicconardi et al. (2013), Pérez-Delgado et al. (2022), Yin et al. (2022)). This situation could result in important species richness sub-estimations within soil arthropod fauna, a substantially misrepresented component of terrestrial animal biodiversity.

DNA barcoding, i.e. the use of short, standardised genomic regions to facilitate species identification and discovery, has revolutionised the study of biodiversity. Barcoding specimens using the standard barcode region (COI-bcr) for metazoan DNA taxonomy (Folmer et al. 1994, Hebert et al. 2003a, Hebert et al. 2003b) can enable the assessment of soil mesofauna diversity in multiple ways. COI-bcr allows the delineation of molecular species, enabling the identification for many animal lineages within morphospecies and the detection of undescribed species (e.g. Janzen et al. (2017), Young et al. (2019)). Similarly, barcoded sequences, through their comparison with available COI-bcr reference databases (e.g. International Barcode of Life Project, iBOL; http://www.ibol.org), provide a global context for species delimitation and identification (e.g. Ashfaq et al. (2017), Cicconardi et al. (2017)). The link between the morphology and the barcode sequences of specimens can be maintained via photographic records or non-destructive DNA extractions and so also provides a fundamental tool for the development of taxonomical knowledge in complex lineages (e.g. Chan et al. (2014), Jackson et al. (2014), Marconi et al. (2022)). Finally, generating COI-bcr databases for specific areas provides a key resource for further implementation of biomonitoring using HTS tools, such as metabarcoding of soil samples (Andújar et al. 2018, Arribas et al. 2019, Andújar et al. 2022). The development of local barcode databases for the biotas of specific countries or regions was started more than a decade ago by pioneering projects (e.g. German Barcode of Life or the Swedish Malaise Trap Project) and it is revealed as a crucial step forward for biodiversity inventory and conservation in those areas (e.g. Hendrich et al. (2015), Young et al. (2019), Ronquist et al. (2020)). Extending these initiatives to understudied biodiversity fractions and hyperdiverse (and threatened) geographical regions is fundamental to the ongoing decline in biodiversity.

The Canary Islands are an oceanic archipelago within the subtropical region of the North Atlantic Ocean with great conservational and patrimonial value in both a national and European context. The Canary Islands are recognised as a Special Territory of the European Union, where quantifying and controlling biodiversity loss is a priority. The diversity of soil mesofauna within oceanic islands needs to be better explored. Literature on the topic is limited (but see Koh et al. (2002), Maraun et al. (2007), Fattorini (2009), Cicconardi et al. (2017)) and even basic species inventory data are, in general, scarce. Within the Canary Islands, the Biodiversity Databank of the Canary Islands (https://

www.biodiversidadcanarias.es/biota/; hereafter referred as BIOTA) is a constantly updated public database containing all species records for the archipelago published in the scientific literature. BIOTA currently reports 474 species of Acari from the Canary Islands, with 288 species from the Island of Tenerife. Recently, a study implementing metabarcoding for a non-exhaustive set of soil samples from Tenerife has revealed nearly double the OTUs of Acari on the Island (Andújar et al. 2022). Remarkably, for most of the lineages (8% similarity clusters from COI metabarcoding data) inventoried in that study, the species-level molecular taxonomic identification was impossible due to the absence of reference barcode sequences.

Here, we initiate the Barcode Database of soil mesofauna from the Canary Islands (CISoilBiota) by: i) developing a standardised workflow that combines traditional morphological identification and COI barcoding of soil arthropod specimens within the framework of the BOLD System (Ratnasingham and Hebert 2007) and ii) by providing the first 168 barcodes from mite specimens collected in soils of the islands of Tenerife and Fuerteventura. This dataset was subsequently analysed as proof of concept to demonstrate the remarkable unrecorded mesofauna diversity present in the soils of the archipelago. Our study aims to provide the basis for the Canary Islands Soil Biodiversity barcode database (CISoilBiota), highlighting its strong potential for the biodiversity inventory and conservation of the soils of the Canary Islands.

Material and methods

Fieldwork and sample processing

Mite specimens were retrieved from 24 soil samples collected between 2018 and 2020 in different localities of laurel forests, pine forests, heathlands and crops of the islands of Tenerife and Fuerteventura (Suppl. material 1). At each site, superficial and deep soil samples were taken. The superficial sample included the litter and the first centimetres of the soil and the deep sample included 25 litres of soil from a hole about 30-40 cm in depth. Every sample was processed following the flotation-Berlese-flotation protocol (FBF) that allows for the 'clean' extraction of arthropod mesofauna (Arribas et al. 2016). Briefly, the FBF protocol is based on soil flotation in water, which allows the extraction of the organic matter and soil mesofauna from raw soil samples. Subsequently, the organic portion is placed in a modified Berlese apparatus to capture specimens alive and preserve them in absolute ethanol at -20ºC. The last step of the FBF protocol is an additional flotation of the ethanol-preserved arthropods, resulting in 'clean' bulk specimen samples. From the 24 cleaned samples of bulk mesofauna, which were examined under a Motic SMZ-171 stereoscope, 239 mites were selected, with representatives from the orders Mesostigmata, Sarcoptiformes and Trombidiformes and maximising the morphological variability of the subset. Each specimen was assigned a unique identifier number (voucher number) and henceforth is referred to as 'vouchers'.

Photo recording and morphological identification

Before DNA extraction, several photos were taken of each intact voucher. A Canon EOS 750D camera attached to a microscope (Zeizz Axioskop 40) was used to take high-resolution pictures of the mite specimens submerged in ethanol and over a white background. For each voucher, a range from 30 to 90 photos was taken by using a 10x objective at different confocal distances depending on the size of each voucher, to subsequently compile with Zerene Stacker (Zerene Systems LLC) for a fully on-focus photographic record of each voucher. Photo alignment and stacking were done following PMax settings and the final image was saved in JPEG format with the highest quality (compression quality = 12). Images and mite vouchers, after non-destructive DNA extractions (see below), were subsequently studied by an expert taxonomist (co-author M.L. Moraza) for morphological identification. Each specimen was examined under a Nikon MSZ745 stereomicroscope and an Olympus Vanox with a phase contrast microscope. For further identification, specimens were cleared in Nesbitt's liquid and mounted using Hoyer's medium. Mite species were identified using taxonomic keys for the Palearctic Region (Gilyarov and Breguetova 1977, Gilyarov 1978, Karg 1993, Pérez-Íñigo 1993, Pérez-Íñigo 1997). Higher mite taxonomic categories follow Krantz and Walter (2009) classification.

DNA extraction and sequencing

Non-destructive DNA extractions were performed for each voucher. For that purpose, the exoskeleton of each specimen was punctured with an entomological pin and kept individually in 1.5 ml vials. Pins were sterilised under a Bunsen burner to avoid contaminations between samples. A volume of 110 µl of digestion buffer (pK ratio 1/10) was added to each voucher and digestion was done overnight at 60°C. The supernatant (DNA lysate) was transferred to the corresponding well within a deep-well plate and the (DNA-extracted) vouchers were maintained in the vials with ethanol for further morphological study (see above). Subsequent steps of the DNA extractions used the MAg-bind Blood & Tissue DNA extraction Kit (Omega Bio-tek) in the KingFisher robotic system (Thermo Fisher Scientific inc.). The default protocol was followed, but DNA lysate and reagent volumes were cut in half. The resulting 100 µl of genomic DNA extraction were split into a 'stock plate' directly frozen at -20ºC and a 'working plate' that was used to quantify DNA concentration using absorbance values within an Infinite M Nano (Tecan Trading AG).

PCR amplification was done for the 5' end COI gene (standard barcode region for Metazoa; Hebert et al. (2003b)) using degenerate Folmer barcode primers (Fol-degen-for: 'TCNACNAAYCAYAARRAYATYGG'; Fol-degen-rev: 'TANACYTCNGGRTGNCCRAA RAAYCA'; Folmer et al. (1994), Yu et al. (2012)). For PCR reaction, 5 µl of extracted DNA was used with 15 µl of PCR mix, which consisted in: 9.72 µl molecular-grade water, 2 µl 10x NH_4 buffer, 1.2 µl $MgCl_2$, 0.4 µl dNTPs, 0.4 µl of BSA, 0.6 µl 10 µM Fol-degen-for and 0.6 µl 10 µM Fol-degen-rev primers and 0.08 µl Taq polymerase (BIOTAQ™ DNA Polymerase, Bioline) per sample. PCR conditions were: 10 min at 95°C in 10 min, followed by 44 cycles of 30 s at 95°C, 30 s at 48°C and 3 min at 72°C; 10 min at 72°C and holding

at 10°C. After viewing PCR products in agarose gel 1%, the positive PCR products were purified. The cleaning mix was prepared with 0.005 μl of exonuclease (exo1), 0.050 μl of rapid alkaline phosphatase (rAP) and 1.945 μl of double distilled water. A volume of 2 μl per 5 μl of DNA sample was used to run a 30 min protocol in the thermal cycler (30 min at 37°C, 5 min at 95°C and holding at 12°C). After cleaning, 5 μl of Fol-degen-for 5 μM primer were added to each sample. The purified PCR product for each voucher was Sanger-sequenced with ABI technology in Macrogen, Spain.

Sequence editing, submission and molecular phylogenetics

Sequences were edited (trimming and primer removal) on Geneious Prime version 2020.0.3 (www.geneious.com). Edited sequences were deposited in BOLD Systems (Ratnasingham and Hebert 2007), together with images, chromatograms and complementary information of the specimens (sampling site, date, taxonomy, store and institution information). Edited sequences were aligned using MAFFT 6.240 (Katoh et al. 2002) with the FFT-NS-i-x2 method and an unweighted pair group method with arithmetic mean (UPGMA) tree from distances corrected under a HKY distance model was generated. In addition, Maximum Likelihood (ML) phylogenetic analyses were run on the IQ-TREE web server at http://iqtree.cibiv.univie.ac.at (Trifinopoulos et al. 2016) using the best fitting substitution model for each codon partition as estimated with ModelFinder (Kalyaanamoorthy et al. 2017). Nodal support was obtained by 1,000 ultrafast boot-strap (UFBoot) replicates (Minh et al. 2013) (Suppl. material 2).

Molecular species delimitation

Five different methods were applied for molecular species delimitation of the vouchers. First, the barcode index number (BIN) was implemented in the BOLD system. This approach provides an effective method for species delineation as each sequence cluster is assigned a unique alphanumeric (BIN URI, see Table 1), which reflects the patterning of intra- and interspecific divergences found in the overall BOLD database (Ratnasingham and Hebert 2007, but see Meier et al. (2022)). Second, the Assemble Species by Automatic Partitioning analysis (ASAP) was implemented. ASAP is a species delimitation method based on pairwise genetic distances ranked using a unique scoring system and consists of merging sequences into groups by an ascending hierarchical clustering until all sequences merge in a single group (Puillandre et al. 2020; https://bioinfo.mnhn.fr/abi/public/asap/). Analysis was implemented over the edited sequences using a K80 Kimura model and the lower ASAP scores as species delimiter parameter (Table 1). Third, we applied the Bayesian implementation of the Poisson tree processes model (PTP, Zhang et al. (2013); http://species.h-its.org/ptp/) by using the phylogenetic tree previously generated on IQ-tree. PTP works with a species delimitation hypothesis, based on the number of mutations (branch lengths) and evaluates species delimitation hypothesis using Maximum-Likelihood algorithms (mlPTP). Finally, as additional criteria for defining molecular entities, we also implemented 3% and 8% genetic distance thresholds using the distance-based UPGMA tree previously generated. The 3% threshold has been widely used to define OTUs in molecular studies (e.g. Hebert et al. (2003a), Magoga et al. (2021)), whereas the

8% threshold has been recently proposed as a conservative threshold in beetles (Salces-Castellano et al. 2021).

Table 1.

Species-level inventory in our dataset, including: morphological identification, number of barcoded sequences, entities delimited for each species delimitation methods (new generated BINs are highlighted by an '*'), representation in BIOTA database (and origin category), similarity percentage with best BOLD match and coherence between species-level identification and best BOLD match identification.

MORPHOLOGICAL IDENTIFICATION	Nº OF SEQUENCES	BIN	ASAP	mlPTP	BIOTA MATCH	NEAREST BOLD MATCH	BOLD COHERENCE
Order Mesostigmata							
Fam. Laelapidae							
Pseudoparasitus dentatus (Halbert, 1920)	2	BOLD:AEI7729*	ASAP21	mlPTP8	yes/native	*Ornithonyssus sylviarum* (79.44%)	no: different species identification
Fam. Macrochelidae							
Macrocheles (*Macrholaspis*) cf. *recki* Bregetova & Koroleva, 1960	1	BOLD:AEI5961*	ASAP54	mlPTP17	yes/native	Mesostigmata (81.9%)	no: higher taxonomic resolution
Fam. Ologamasidae							
Gamasiphis sextus Vitzthum, 1921	4	BOLD:AEI1055*	ASAP13	mlPTP18	no	*Gamasiphis* sp. JCS03 (83.72%)	no: higher taxonomic resolution
Fam. Parasitidae							
Holoparasitus sp.	3	BOLD:AEI8993*	ASAP6	mlPTP23	NA	Parasitidae (83.00%)	NA
Parasitidae	1	BOLD:AEI6503*	ASAP39	mlPTP22	NA	*Poecilochirus* (84.57%)	NA
Pergamasus crassipes (Linnaeus, 1758)	1	BOLD:ACQ8500	ASAP2	mlPTP21	yes/introduced	Mesostigmata (98.34%)	no: higher taxonomic resolution
Fam. Polyaspididae							
Uroseius cylindricus (Berlese, 1916)	1	BOLD:AEI1534*	ASAP44	mlPTP45	no	*Polyaspinus higginsi* (84.23%)	no: different species identification
Fam. Trachyuropodidae							

MORPHOLOGICAL IDENTIFICATION	Nº OF SEQUENCES	BIN	ASAP	mlPTP	BIOTA MATCH	NEAREST BOLD MATCH	BOLD COHERENCE
cf. *Trachyuropoda* sp.	1	BOLD:ADH9050	ASAP35	mlPTP43	NA	Uropodidae (99.5%)	NA
Trachyuropoda sp.	5	BOLD:AAZ2213	ASAP1	mlPTP44	NA	Mesostigmata (98.54%)	NA
Order Trombidiformes							
Fam. Erythraeidae							
cf. *Leptus* sp.	2	BOLD:AEI0533*	ASAP47	mlPTP66; mlPTP67	NA	Arachnida (88.08%)	NA
Erythraeidae	1	BOLD:AEI6506*	ASAP38	mlPTP7	NA	Trombiculidae (78.51%)	NA
Order Sarcoptiformes							
Fam. Achipteriidae							
Campachipteria petiti (Travé, 1960)	3	BOLD:AEI0246*; BOLD:AEI6330*	ASAP20	mlPTP54; mlPTP59; mlPTP60	no	Achipteriidae (86.7%)	no: higher taxonomic resolution
Fam. Ameridae							
Amerus cuspidatus (Berlese, 1883)	2	BOLD:AEI0247*; BOLD:AEI4260*	ASAP23	mlPTP57; mlPTP58	yes/native	*Ceratozetes gracilis* (83.02%)	no: different species identification
Fam. Amerobelbidae							
Amerobelba decedens Berlese, 1908	3	BOLD:AEI8991*	ASAP7	mlPTP38	yes/native	*Eueremaeus* (83.33%)	no: higher taxonomic resolution
Fam. Carabodidae							
Carabodidae	1	BOLD:AEI9656*	ASAP40	mlPTP29	NA	Hermanniellidae (80.95%)	NA
Cavernocarabodes trigonosternum (Pérez-Íñigo, 1976)	1	BOLD:AEI0245*	ASAP53	mlPTP37	yes/ endemic	*Oribatodes mirabilis* (85.96%)	no: different species identification
Odontocepheus elongatus (Michael, 1879)	1	BOLD:AEI0250*	ASAP60	mlPTP2	yes/native	*Odontocepheus elongatus* (74.48%)	yes
Fam. Ceratoppiidae							

MORPHOLOGICAL IDENTIFICATION	Nº OF SEQUENCES	BIN	ASAP	mlPTP	BIOTA MATCH	NEAREST BOLD MATCH	BOLD COHERENCE
Ceratoppia bipilis (Hermann, 1804)	1	BOLD:AEH9721*	ASAP49	mlPTP34	yes/native	*Ceratoppia* (82.96%)	no: higher taxonomic resolution
Fam. Ceratozetidae							
Trichoribates novus (Sellnick, 1928)	2	BOLD:AEI6507*	ASAP41	mlPTP46	no	*Oribatella* (84.86%)	no: higher taxonomic resolution
Fam. Compactozetidae							
Cepheus latus Koch, 1835	1	BOLD:AEI5729*	ASAP55	mlPTP36	yes/native	Neoliodidae (82.41%)	no: higher taxonomic resolution
Conoppia cf. *palmicincta* (Michael, 1884)	1	BOLD:AEI8929*	ASAP46	mlPTP48	yes/native	*Eremaeus* (88.68%)	no: higher taxonomic resolution
Fam. Damaeidae							
Damaeus recasensi Capilla, 1971	10	BOLD:AEI4384*; BOLD:AEI4385*; BOLD:AEI4386*; BOLD:AEI7071*	ASAP14	mlPTP70; mlPTP71	yes/native	*Epidamaeus* (83.67%)	no: higher taxonomic resolution
Metabelbella interlamellaris Pérez-Íñigo, 1987	2	BOLD:AEI0249*	ASAP45	mlPTP24	yes/native	Damaeidae (80.81%)	no: higher taxonomic resolution
Fam. Dampfiellidae							
Dampfiella ambigua Pérez-Íñigo, 1976	1	BOLD:AEI5421*	ASAP50	mlPTP1	yes/ endemic	*Baryscapus servadeii* (77.57%)	no: different species identification
Fam. Euphthiracaridae							
Acrotritia ardua ardua (Koch, 1841)	2	BOLD:AAF9157	ASAP57	mlPTP33	yes/native	Euphthiracaridae (95.34%)	no: higher taxonomic resolution
Acrotritia penicillata (Pérez-Íñigo, 1969)	1	BOLD:ADX1060	ASAP30	mlPTP32	no	Sarcoptiformes (99.5%)	no: higher taxonomic resolution
cf. *Euphthiracarus* sp.	1	BOLD:AEH9008*	ASAP56	mlPTP14	NA	*Euphthiracarus monodactylus* (81.03%)	NA

MORPHOLOGICAL IDENTIFICATION	Nº OF SEQUENCES	BIN	ASAP	mlPTP	BIOTA MATCH	NEAREST BOLD MATCH	BOLD COHERENCE
Mesotritia cf. *grandjeani* (Feider & Suciu, 1957)	9	BOLD:AEI8467*; BOLD:AEI6713*	ASAP22	mlPTP3	no	Arthropoda (78.86%)	no: higher taxonomic resolution
Fam. Galumnidae							
Acrogalumna longipluma (Berlese, 1904)	19	BOLD:AEI1056*; BOLD:AEI5290*	ASAP9; ASAP10	mlPTP55; mlPTP56	yes/native	*Acrogalumna longipluma* (96.45%)	yes
Galumna alata (Hermann, 1804)	4	BOLD:AEI4158*; BOLD:AEI8205*	ASAP17; ASAP18	mlPTP49; mlPTP68; mlPTP69	yes/native	*Eupelops* (82.09%)	no: higher taxonomic resolution
Pilogalumna allifera (Oudemans, 1919)	2	BOLD:AEH9722*	ASAP31	mlPTP19	yes/ endemic	*Cepheus* (82.72%)	no: higher taxonomic resolution
Fam. Gustaviidae							
Gustavia longirostris Mihelcic, 1957	4	BOLD:AEI3725*	ASAP33	mlPTP26	no	*Chamobates cuspidatus* (81.92%)	no: different species identification
Fam. Humerobatidae							
Humerobates pomboi Pérez-Íñigo, 1992	1	BOLD:AEI4261*	ASAP29	mlPTP47	yes/native	Humerobatidae (86.53%)	no: higher taxonomic resolution
Fam. Hypochthoniidae							
Hypochthonius luteus Oudemans, 1917	18	BOLD:AEI3587*	ASAP5	mlPTP13	yes/native	*Hypochthonius luteus* (96.34%)	yes
Fam. Liacaridae							
Dorycranosus splendens (Coggi, 1898)	1	BOLD:AEI6712*	ASAP26	mlPTP4	yes/native	Oppiidae (76.08%)	no: higher taxonomic resolution
Fam. Nothridae							
Nothrus reticulatus Sitnikova, 1975	1	BOLD:AEI6500*	ASAP43	mlPTP39	no	*Nothrus* (95.27%)	no: higher taxonomic resolution
Nothrus silvestris Nicolet, 1855	2	BOLD:AEI0848*	ASAP28	mlPTP40	yes/native	*Nothrus* (82.57%)	no: higher taxonomic resolution

MORPHOLOGICAL IDENTIFICATION	Nº OF SEQUENCES	BIN	ASAP	mlPTP	BIOTA MATCH	NEAREST BOLD MATCH	BOLD COHERENCE
Fam. Oppiidae							
Ramusella cf. *clavipectinata* (Michael, 1885)	1	BOLD:AEI7727*	ASAP32	mlPTP28	yes/native	*Eremaeus* (84.67%)	no: higher taxonomic resolution
Fam. Oribatulidae							
Hemileius elongatus E.Pérez-Íñigo, 1978	4	BOLD:AEI4159*	ASAP8	mlPTP20	yes/native	*Hemileius initialis* (85.82%)	no: different species identification
Zygoribatula connexa (Berlese, 1904)	1	BOLD:AEI7730*	ASAP16	mlPTP25	yes/native	*Oribatula tibialis* (85.02%)	no: different species identification
Zygoribatula propinqua (Oudemans, 1902)	2	BOLD:AEI7725*; BOLD:AEI8260*	ASAP34; ASAP59	mlPTP27	yes/native	*Eueremaeus silvestris* (84.85%)	no: different species identification
Zygoribatula undulata Berlese, 1916	2	BOLD:AEI7728*	ASAP12	mlPTP31	yes/native	*Achipteria coleoptrata* (86.16%)	no: different species identification
Fam. Phenopelopidae							
Eupelops acromios (Hermann, 1804)	5	BOLD:AEI2704*; BOLD:AEI3535*	ASAP36	mlPTP64; mlPTP65	yes/native	*Eupelops* (84.57%)	no: higher taxonomic resolution
Fam. Phthiracaridae							
Archiphthiracarus sp.	2	BOLD:AEI0248*; BOLD:AEI7726*	ASAP42; ASAP58	mlPTP10; mlPTP16	NA	*Phthiracarus globosus* (95.78%)	NA
Hoplophthiracarus cf. *cazanicus* Feider & Calugar, 1970	1	BOLD:AEH8933*	ASAP24	mlPTP15	no	*Austrophthiracarus costai* (80.35%)	no: different species identification
Phthiracarus cf. *globosus* (Koch, 1841)	13	BOLD:AEI2040*	ASAP11	mlPTP11	no	*Phthiracarus* (79.94%)	no: higher taxonomic resolution
Phthiracarus cf. *globus* Parry, 1979	1	BOLD:AEI5730	ASAP25	mlPTP9	no	*Phthiracarus* (98.04%)	no: higher taxonomic resolution
Phthiracarus cf. *ligneus* Willmann, 1931	5	BOLD:AEH9005*; BOLD:AEI6501*; BOLD:AEI7167*	ASAP27	mlPTP61; mlPTP62; mlPTP63	no	*Phthiracarus globosus* (78.43%)	no: different species identification

MORPHOLOGICAL IDENTIFICATION	Nº OF SEQUENCES	BIN	ASAP	mIPTP	BIOTA MATCH	NEAREST BOLD MATCH	BOLD COHERENCE
Fam. Steganacaridae							
Steganacarus tenerifensis Pérez-Íñigo, 1972	2	BOLD:AEI8994*	ASAP19	mIPTP12	yes/ endemic	*Steganacarus magnus* (84.21%)	no: different species identification
Fam. Suctobelbidae							
Rhynchobelba machadoi Pérez-Íñigo, 1976	1	BOLD:AEI6502*	ASAP48	mIPTP6	yes/ endemic	*Neogymnobates luteus* (81.77%)	no: different species identification
Fam. Tectocepheidae							
Tectocepheus alatus Berlese, 1913	1	BOLD:AEI6504*	ASAP37	mIPTP5	no	*Scutovertex sculptus* (85.44%)	no: different species identification
Fam. Trhypochthoniidae							
Trhypochthonius japonicus Aoki, 1970	1	BOLD:AEI0244*	ASAP52	mIPTP30	no	*Trhypochthonius tectorum* (82.95%)	no: different species identification
Fam. Xenillidae							
Xenillus discrepans canariensis Pérez-Íñigo, 1976	1	BOLD:AEH9009*	ASAP51	mIPTP42	yes/ endemic	Arachnida (80.16%)	no: higher taxonomic resolution
Xenillus sp1	1	BOLD:AEI9187*	ASAP15	mIPTP50	NA	Sarcoptiformes (81.48%)	NA
Xenillus sp2	1	BOLD:AEH9719*	ASAP3	mIPTP41	NA	Scheloribatidae (78.5%)	NA
Xenillus sp3	2	BOLD:AEI3741*	ASAP4	mIPTP52	NA	*Parachipteria punctata* (80.93%)	NA
Xenillus sp4	3	BOLD:AEI6505*	ASAP4	mIPTP53	NA	Sarcoptiformes (80.81%)	NA
Xenillus tegeocranus (Hermann, 1804)	2	BOLD:AEI8992*	ASAP15	mIPTP51	yes/native	Liacaridae (82.13%)	no: higher taxonomic resolution

Evaluating unrecorded and cryptic diversity within the Canary Islands context

To evaluate the previously unrecorded diversity in the Canary Islands that our dataset contains, we performed searches of the different species identified in our data within BIOTA. For each species within our dataset, we annotated the previously existing records and their status as endemic, native or potentially introduced taxa in the archipelago context.

To identify potential cryptic diversity within our dataset, we compare the morphological assignment provided to specimens with the results provided by the different species delimitation methods implemented. The number of morphological species that were split, merged or maintained was estimated for each delimitation method and the concordance amongst the specimens grouping resulting from each approach was evaluated.

Evaluating unrecorded and cryptic diversity in the global context

To quantify the overall unrecorded mite diversity within our dataset, we used the Barcode of Life Data (BOLD) as the major source of barcode reference sequences available. BOLD contains 17,789,385 specimen records, of which 13,911,307 have barcode sequences (accessed at 02/08/2023). Of these, a total of 219,759 records are from mites, of which 181,682 have barcode sequences. Records with barcode sequence and identification to species level represent a total of 4,350 mite species, included in orders Holothyrida (1), Ixodida (334), Mesostigmata (992), Sarcoptiformes (769) and Trombidiformes (2,254). We have compared our sequences with those available on the platform to check the consistency of morphological identifications, detect potential cryptic diversity at the global scale and to investigate to what extent the diversity within our dataset is already represented in the BOLD repository.

First, we performed BOLD Identification System (IDS) (default setting parameters) for each sequence against the overall BOLD system (21/04/2023), including public and private barcode records. We extracted the taxonomic identification, based on morphology (species or genus level) and the similarity percentage to the best match using the BOLD Identification System (IDS) results for each sequence. Using this information: i) we estimated the overall similarity with BOLD sequences for each of our species and ii) we checked the coherence between the taxonomic identifications from BOLD and the morphological identifications of our specimens. In the cases where species-level identification agreed between both datasets, the overall similarity, the monophyly of the Canary Islands sequences and the geographical origin of BOLD sequences were evaluated to identify potential cryptic diversity within those mite species. Finally, as an additional indicator of how the diversity in our dataset is already reported within BOLD, the number of BINs that were only composed by our sequences (i.e. not including sequences already present on the repository) was recorded.

Results

From the total 239 mite specimens selected, we recovered 168 barcode sequences, resulting in a success rate above 70%. Sequence lengths varied from 472 to 639 bp and included 153 haplotypes. The 168 barcoded specimens were morphologically identified as corresponding with 58 different morphological entities, of which 45 correspond to known species, four to unknown, likely new species of the genus *Xenillus* and nine entities (each represented by a single specimen) that could only be identified at the genus (six cases) or family (three cases) levels. Our barcoded dataset comprises entities from three orders (Mesostigmata, Sarcoptiformes and Trombidiformes) and 32 families, including Xenillidae (six entities), Phthiracaridae (five entities) and Euphthiracaridae and Oribatulidae (four entities each) as the four families with a higher number of species in our dataset (Table 1, Suppl. material 3). All sequences and associated metadata were included in the BOLD System within the project CISoilBiota, subproject CIACA (Acari of the Canary Islands) as part of the BOLD Campaign 'Fauna of the Canary Islands'. Samples were named for the submission to BOLD Systems as CIACA001-21 to CIACA168-21 (see Suppl. material 4). See Figs 1, 2 for a schematic representation of the workflow and a subset of the photographic records uploaded to BOLD.
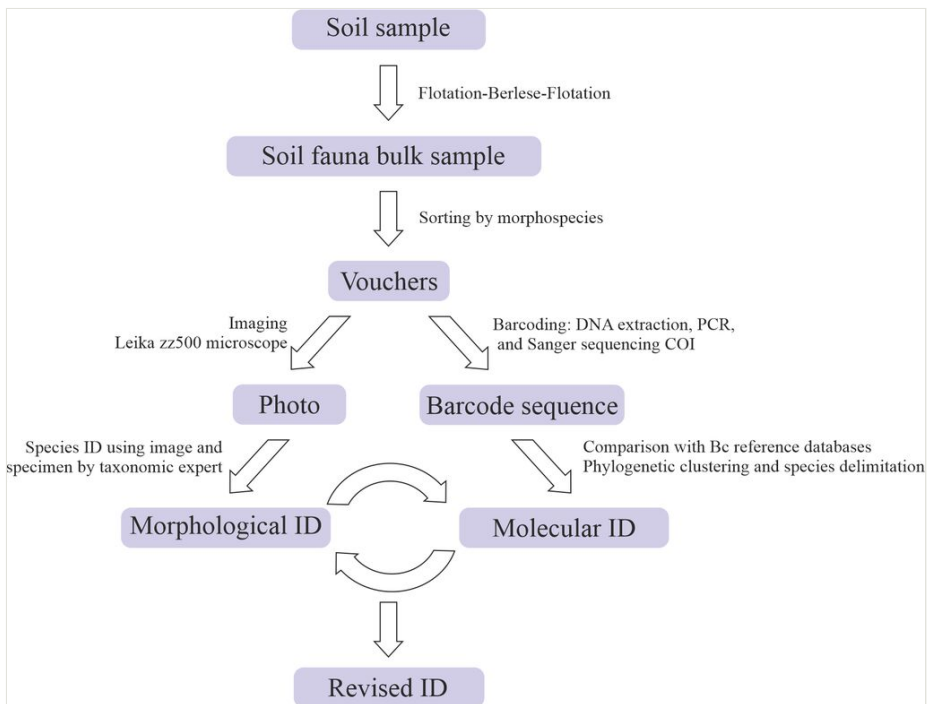

Figure 1. doi

Schematic representation of the implemented workflow that combines traditional morphological identification and COI barcoding of soil arthropod specimens.
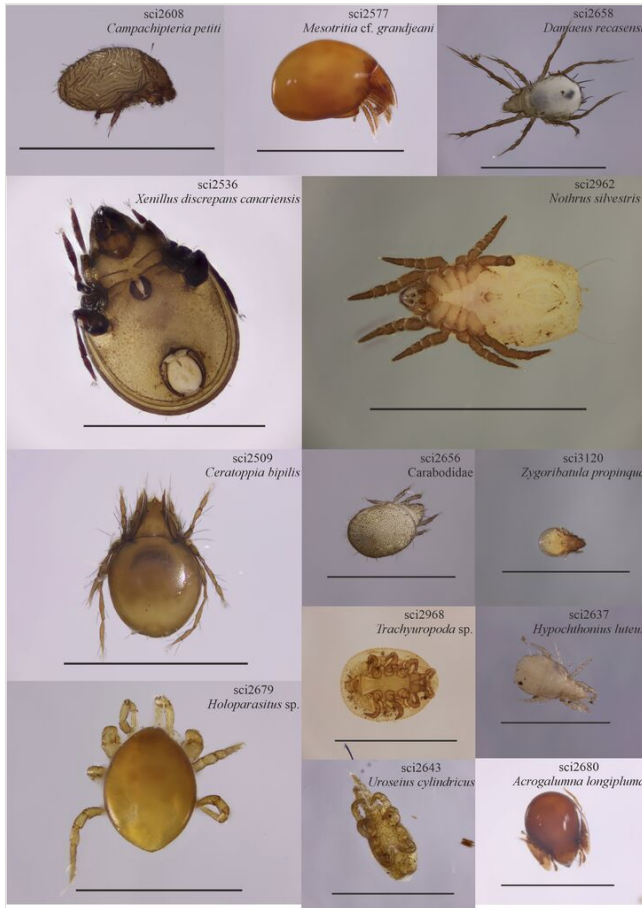
Figure 2. doi

Representative examples of obtained images from voucher specimens. Scale = 1 mm. For specimen codes, see Suppl. materials 3, 4.

Morphological and molecular species delimitations: evaluating cryptic diversity within the Canary Islands context

The different molecular species delimitation methods implemented resulted in a range from 60 to 85 molecular species delimited. The barcode index number (BIN) implemented in the BOLD System showed that our 168 voucher sequences were grouped in 71 BINs. Sixty-five of those BINs were newly generated by BOLD and assigned to sequences contributed in this study, while 11 sequences were grouped into six pre-existing BINs (access numbers in Table 1). ASAP analysis grouped our sequences in 60 molecular species according to the lower ASAP score, whereas PTP analysis delimited 71 molecular species for the mlPTP approach. Finally, the 3% genetic distance threshold resulted in 70 groups and the 8% genetic distance threshold in 64 (Figs 3, 4).
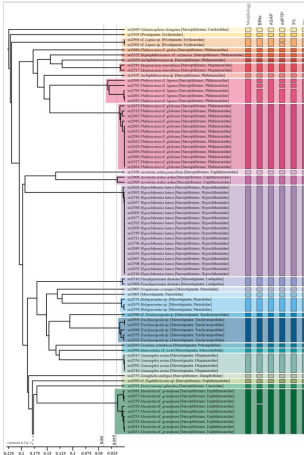
Figure 3. doi

Distance-based UPGMA tree obtained using HKY corrected distances. Coloured horizontal blocks over the tree represent specimen clusters corresponding to morphological species. Vertical bars represent, from left to right: (i) morphological species, (ii) BINs classification in BOLD, (iii) species delimitation with ASAP, (iv) species delimitation with mlPTP, (v) 3% similarity clusters and (vi) 8% similarity clusters. At the bottom, each method's total number of species is presented. The X-axis represents genetic distance; with dotted lines corresponding 3% and 8% divergence thresholds (first half).

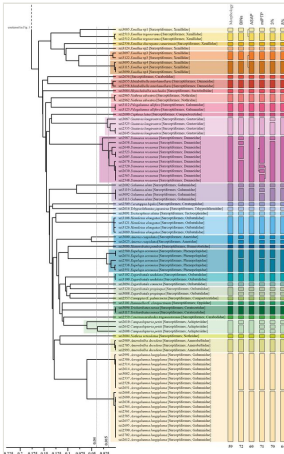Figure 4. doi

Distance-based UPGMA tree obtained using HKY corrected distances. Coloured horizontal blocks over the tree represent specimen clusters corresponding to morphological species. Vertical bars represent, from left to right: (i) morphological species, (ii) BINs classification in BOLD, (iii) species delimitation with ASAP, (iv) species delimitation with mlPTP, (v) 3% similarity clusters and (vi) 8% similarity clusters. At the bottom, each method's total number of species is presented. The X-axis represents genetic distance; with dotted lines corresponding 3% and 8% divergence thresholds (second half).

The concordance amongst the results of the different molecular delimitations implemented was relatively high. BINs, mPTP and 3% approaches resulted in a higher (but mostly concordant) number of delimited entities (Figs 3, 4). In all cases, molecular species delimitations exceeded the 58 entities morphologically delimited. In general, morphologically identified species and molecular species groupings were coincident, but at least eleven morphologically delimited species consistently included multiple molecular entities in at least one delimitation method (*Phthiracarus* cf. *ligneus*, *Mesotritia* cf. *grandjeani*, *Gustavia longirostris*, *Damaeus recasensi*, *Galumna alata*, *Amerus cuspidatus*, *Eupelops acromios*, *Zygoribatula propinqua*, *Campachipteria petiti*, and *Acrogalumna longipluma*) (Figs 3, 4).

Comparison with the Biodiversity Databank of the Canary Islands: evaluating unrecorded diversity within the Canary Island context

Of the 45 morphological entities identified as known species in our dataset, 31 species (68%) were already registered as present in the archipelago, whereas the remaining 14 (31%) represented new species and genera records at the archipelago level (Table 1, Fig. 5a).
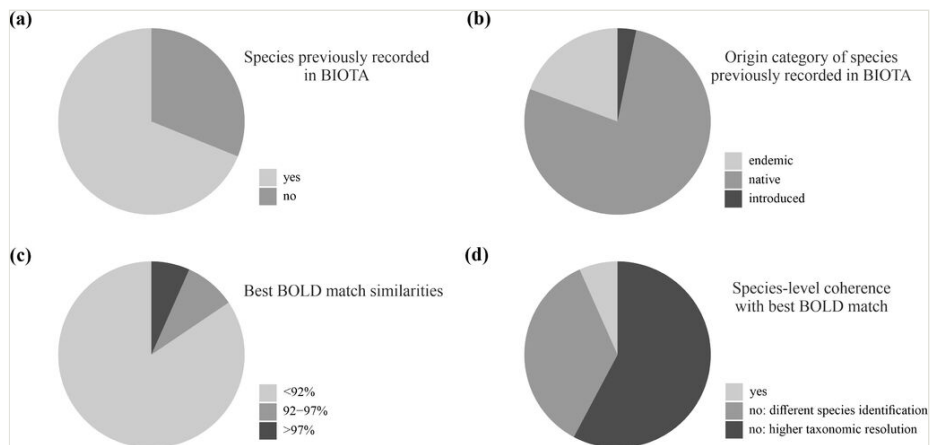

Figure 5. doi

Evaluating unrecorded diversity within the Canary Islands and global context. **a** Proportion of species in our dataset that are already registered in the BIOTA database; **b** Category of origin (i.e. native non-endemic, introduced or endemic) for species recorded in BIOTA as reported within the database; **c** Similarity values of best matches of obtained sequences representing each species against BOLD Systems; **d** Species-level identification coherence between the specimens in our dataset and BOLD best matches. When there is no coherence, we specify if the BOLD best match was identified at species or higher taxonomic level (genus or family level).

Of the 31 species in our dataset that are present in the BIOTA, six species are endemic to the Islands (19%) and the rest are considered non-endemic native species (24 species, 77%) or introduced species (one species, 3%) (Table 1, Fig. 5b).

Comparison with the BOLD system: Evaluating unrecorded and cryptic diversity in the global context

BOLD Identification System (IDS) results for the 168 sequences in our dataset resulted in a range of similarity percentages, with the best matches ranging from 74.48% to 99.5%. Amongst the 45 morphological entities identified as known species, 38 (84%) reported a similarity below 92%, four (9%) from 92% to 97% and three (7%) above 97% (Table 1, Fig. 5c). Regarding the taxonomic identity of the BOLD best matches for our sequences, the best match has no species-level identification for 26 (58%) entities. For the remaining cases where best BOLD matches have species-level identifications, three (6%) are concordant with the morphological identifications of our specimens and 16 (35%) showed no concordant species-level identifications (Table 1, Fig. 5d).

The only cases where species-level identification agrees with the BOLD dataset are *Acrogalumna longipluma*, *Hypochthonius luteus* and *Odontocepheus elongatus*. In the case of *Acrogalumna longipluma*, the overall similarity of Canarian specimens with BOLD sequences was 96.45%. In the case of *Hypochthonius luteus*, the overall similarity of our Canarian specimens with the only one registered on BOLD was 96.34%. Finally, in the case of *Odontocepheus elongatus*, the overall similarity of Canarian specimens with BOLD sequences was 74.48%.

Discussion

A molecular barcoding framework for the soil fauna of the Canary Islands

Our study initiates the Barcode Database of soil fauna from the Canary Islands (CISoilBiota) by developing a standardised workflow that combines specific soil sampling, Berlese extraction, sample sorting, COI barcoding and traditional taxonomic identification of barcoded specimens. The workflow has been applied to 239 mite specimens, of which we recovered 168 sequences. This represents a success of 70%, similar to success rates in other barcoding studies (e.g. deWaard (2019), Salces-Castellano et al. (2021), Suárez et al. (2022), Suárez et al. (2023), Caterino and Recuero (2023)). We did not detect any pattern amongst failures regarding taxonomical assignments or geographic distribution of soil samples. Causes of failures may be indicative of a low quantity of DNA retrieved, considering that we worked with minute mesofauna specimens or poor quality of DNA, as specimens were collected using Berlese apparatus with water on the collecting recipients. Each Berlese apparatus was revised and specimens were transferred to ethanol every two days to minimise the degradation of DNA. Still, we cannot discard DNA degradation as affecting PCR performance in some cases. Although extracted DNA quantification and further dilutions or reconcentration will help to obtain a higher success rate, we consider this 70% success rate as a good starting point for further development, considering inherent difficulties of DNA work with small-sized soil mesofauna.

One of these difficulties is the incompatibility of the procedures used for the morphological study of these minute organisms (requiring microscopic preparations where specimens are

cleared and fixed with different chemical products) and DNA preservation. Here, we solve that by implementing a protocol of specimen imaging and non-destructive DNA extractions for mites that allow the morphological study of the specimens after DNA extraction. Our results demonstrated that non-destructive DNA extraction of soil mites is feasible without compromising the morphological integrity of specimens.

Another difficulty in implementing barcoding to soil mesofauna is associated with the reduced body size of specimens and the low DNA concentration retrieved. The DNA extraction and PCR protocol performed here appears adequate under these low DNA conditions. DNA extraction was implemented using a magnetic-bead approach in a robotic platform; this semi-automated approach is optimal for implementing arthropod barcoding because it facilitates the standardised processing of high numbers of specimens while maximising the quality of DNA extracts for long-term storage (Arribas et al. 2022). We reduced reagent volumes in half without an evident impact on DNA extraction performance. Further tests with more reduced volumes would be desirable to minimise costs associated with DNA extraction. Finally, the high phylogenetic diversity within soil mites could challenge the selection of the primer sets for PCR amplification of the barcode fragment. Our results, aligned with previous studies (Young et al. 2012, Arribas et al. 2016, Arribas et al. 2019), demonstrate an overall good performance of the Folmer degenerate barcode primers (Yu et al. 2012) for the broad diversity of soil mites.

We expect that barcoding effort over soil mites can be additionally improved by the application of High Throughput Sequencing (HTS) approaches, at the same time that costs are reduced (Srivathsan et al. 2018, Creedy et al. 2021, Srivathsan et al. 2021, Emerson et al. 2022). Although Sanger sequencing is an efficient and optimised technique, multiplexing approaches combined with the strength of HTS can be used to generate thousands of barcode sequences in a faster and cheaper way (Srivathsan et al. 2019, Srivathsan et al. 2023, Vasilita et al. 2023). The rigorous implementation of HTS barcoding methodologies on the remarkably abundant and hyperdiverse soil mesofauna of arthropods holds great promise in addressing the global lack of knowledge on soil biodiversity. Still, to maximise the utility of obtained sequences as barcode references, protocols as proposed here, are fundamental, as the link amongst obtained sequences, properly preserved voucher specimens and images is maintained. These procedures include costly and time-consuming steps for specimen sorting, imaging and puncturing before DNA extraction, in addition to requiring great taxonomic expertise. We acknowledge that morphological identification is the main bottle-neck for the whole approach and, in agreement with Srivathsan et al. (2021), we envision a system where HTS barcoding can be used to obtain barcode sequences from high numbers of soil mesofauna specimens that, after the application of DNA similarity clustering and molecular identification methods, can be subsampled for a detailed morphological study of a much-reduced number of representative specimens.

Barcoding to unveil cryptic diversity in soil fauna

The molecular species delimitations showed broad consistency amongst them and with morphological species identifications in our dataset (Figs 3, 4). ASAP and 8% threshold showed a higher agreement in delimited entities between them and with the morphological identification, whereas BINs, mlPTP and 3% resulted in additional splits for some lineages, likely representing over-splitting as previously reported in other studies (Copilaş-Ciocianu et al. 2022, Ranasinghe et al. 2022). Still, the comparison of the more conservative ASAP and 8% threshold with the morphological identification allowed us to detect several cases of inconsistency and potential cryptic diversity. The taxonomic challenge posed by cryptic species (two or more morphologically similar species classified as a single species) has been recognised as a significant limitation in quantifying soil mesofauna (Emerson et al. 2011, Pfingstl et al. 2021, Szudarek-Trepto et al. 2021, Yin et al. 2022). Different studies, implementing DNA sequencing, have revealed that cryptic diversity could be massive in specific soil lineages (Cicconardi et al. 2010, Cicconardi et al. 2013). In this study, despite our reduced sampling (168 barcodes from mite specimens), we evidenced several examples of potential cryptic diversity within the mites of the Canary Islands, where molecular delimitations identify multiple divergent lineages within a single morphological species.

We have detected a series of cases where specimens, morphologically identified as a species, show intraspecific divergences higher than 3%. Part of these cases consists of monophyletic lineages where internal divergences are higher than 3%, but lower than 8%. Here we found the cases of: (i) *Phthiracarus* cf. *ligneus* with three lineages with divergences over 6%; (ii) *Damaeus recasensi* with two lineages (one with a single specimen) with divergences above 5%; (iii) *Eupelops acromios* with two lineages with divergences above 3%; and (iv) *Gustavia longirostris* with two lineages with divergences above 3%. In these four cases, the moderately high intraspecific divergences found are compatible with a single species, which is also suggested by molecular species delimitation methods, such as ASAP (Figs 3, 4). In fact, high intraspecific variation can be expected for soil mites if we consider the huge population sizes reported for some soil taxa (Petersen and Luxton 1982, Endlweber et al. 2006) and the expectations from the neutral theory of molecular evolution, with genetic diversity increasing with a larger effective population size and the decreasing effects of drift (Kimura 1979). Still, these divergences also suggest that additional attention should be placed on these lineages, as they likely represent native species within the Canaries, which is a highly fragmented landscape that can contribute to geographic isolation and diversification (Juan et al. 2000). Alternatively, this pattern may also reflect the human-mediated introduction of new populations for Canarian native species or even cases of multiple introduction of non-native species. More detailed studies will be required to distinguish amongst the different alternatives for each lineage.

We have also detected other cases where specimens classified as a single species are split into different lineages with divergences higher than 8% for the barcode fragment (Figs 3, 4). These cases include: (i) *Campachipteria petiti*, represented by three specimens with

divergences between 7.1 and 11.1%; (ii) *Galumna alata* with two lineages diverging 14.3%; (iii) *Amerus cuspidatus* represented by only two specimens having a divergence of 9.3%; (iv) *Acrogalumna longipluma* with two lineages with divergences of 11.5%; (v) *Zygoribatula propinqua* represented by only two specimens having a divergence of 17.5%; (vi) *Acrotritia ardua*, where the two morphological subspecies show molecular divergences of 21.7% and (vii) the two morphologically highly similar specimens classified as *Archiphthiracarus* sp., which are only distantly related phylogenetically (26.9% divergence). All these cases suggest speciation with reduced morphological differentiation, although deep mitochondrial DNA divergence has been also shown to not indicate distinct species in some lineages (e.g. Leo et al. (2010)). A detailed morphological study and the sequencing of additional genes from the nuclear genome and specimens from a wider geographical range will be needed to clarify their taxonomic status.

Further implementation of the proposed barcoding workflow within the Canary Islands will contribute to elucidating the status of the reported cases and, presumably, to detect additional cases of cryptic diversity. An integrative approach, with parallel and interactive morphological and molecular work, will contribute to accelerating species inventory and discovery. For example, in our dataset and within the genus *Xenillus*, two already-described species are detected, with additional specimens showing morphological variation not matching any described species. Molecular analysis shows consistent results with morphology, suggesting the existence of four additional new species within the genus with divergences above 13.2% (Figs 3, 4).

Beyond the prevalence of cryptic diversity within the soil mites of the Canary Islands, our results point to the generality of this pattern globally. The analyses comparing our sequences with the BOLD database found three paradigmatic cases of potential cryptic diversity within worldwide distributed species. The first one is the case of *Acrogalumna longipluma*, with available barcode sequences from Canada, Germany, Finland, UK and the Canary Islands, forming five differentiated geographically coherent lineages with similarities below 97%. Of these, two lineages are exclusively found in the Canary Islands with divergences higher than 12% and not showing a sister taxa relationship (Suppl. materials 5, 6). This pattern suggests that, under the taxonomic name *Acrogalumna longipluma*, we currently enclose a complex of species, that could include two endemic species for the Canary Islands and that requires a detailed morphological revision. The second case is that of *Hypochthonius luteus*, where the Canarian specimens form a lineage sister to a specimen from Belgium, with an overall similarity of 96.34%, again pointing to potential cryptic diversity within this morphologically described species (Suppl. material 7). In the third case, that of *Odontocepheus elongatus*, available sequences from Norway, Finland and the Canary Islands form three differentiated lineages with similarities below 75%, one of these lineages exclusively composed of Canarian specimens, showing again cryptic diversity within a highly similar morphology (Suppl. material 8).

Barcoding to unveil the (unknown) diversity of soil fauna within the Canary Islands

If a reference database is poorly populated for a specific group, the probability of inacurate taxonomic assignment is higher and placement to high taxonomic ranks is frequent (Somervuo et al. 2017). Our results illustrate the high magnitude of unrecorded diversity within soil mites and associated issues when comparing obtained barcodes with reference databases. Most of the taxa within our dataset were not previously represented in the BOLD database, with most generated barcode sequences showing low similarity values with the best available matches in BOLD. In addition, a high proportion of best matches corresponded to specimens without a species-level identification in BOLD (26 cases) or to different species to those here identified morphologically (16 cases, all with low similarity). The lack of morphological identifications to the species level can be understood given: (i) the great species richness, abundance and local-scale heterogeneity of mesofauna communities, (ii) the poor taxonomic background knowledge of many soil lineages and geographic areas and (iii) the general lack of taxonomic expertise on soil mites (André et al. 1994, Decaëns 2010, Cameron et al. 2018, Guerra et al. 2020, White et al. 2020). Consequently, species widely distributed, naturally or by mean of human introductions, are likely better represented in public repositories. In contrast, those species locally endemic with small distributional ranges are represented to a lesser extent and only from a few geographical regions (Porter and Hajibabaei 2018). This lack of representation and taxonomic resolution in reference databases limits their potential to provide a reliable taxonomic assignment to newly-generated barcodes, highlighting the imperative need for further barcode projects on soil fauna integrating taxonomic expertise.

Within the context of the Canary Islands, our results point to a massive under-representation of the diversity of soil mites in biodiversity databases. The BIOTA database contains records for 474 species and subspecies of mites, of which 425 are considered native species and 49 introduced species. Of those species classified as native, 110 species are considered endemic to the Canary Islands and 104 species endemic to the Macaronesian Region. Regarding our data, 14 of the 45 (31%) species, for which we obtained a species-level identification, represent the first record for the Canary Islands, all of them also providing the first record at the genus-level. All of these are species known from outside the Canaries and are now reported to the Canaries for the first time. These species may correspond to native non-endemic species or introduced species, according to their known distribution outside the Canaries, but given the absence of reference sequences for most of these species, we cannot discard that they represent additional cases of cryptic diversity. For example, in our dataset, we found four cases in which, although it is not the best match in BOLD, there has been a match with a sequence of the same species. *Acrotritia ardua ardua* and *A. penicillata* have a 78.86-79.57% similarity with several sequences named *Rhysotritia ardua* (junior synonym) from Canada, Poland and Norway (see Suppl. materials 9, 10). *Phthiracarus* cf. *globosus* has a 77.78-77.89% similarity with sequences from Finland (see Suppl. material 11). *Xenillus tegeocranus* has a 78.15-80.07% similarity with sequences from Canada, Finland, Norway and Slovakia (see Suppl. material 12). These worldwide matches suggest that species that have been

taxonomically assigned to a single species may constitute cryptic lineages. To advance our knowledge of the magnitude of biodiversity and level of endemicity of the Canary Islands, we will need further efforts in generating barcode reference sequences, with reliable taxonomic identification, from inside the Canaries, but also from other regions. Barcoding specimens unambiguously associated with a particular species considering both its morphology and geographic origin is needed. These data will be key to distinguishing amongst native, endemic and introduced species within the Canaries, allowing us to generate reliable local inventories of soil fauna.

Conclusions

This study provides and demonstrates the efficiency of a standardised workflow that combines traditional morphological identification and COI barcoding for the challenging soil fauna of mites within the framework of the BOLD System. Despite our reduced sampling, our results on interrogating the generated biodiversity data demonstrate the remarkable unrecorded mesofauna diversity present in the soils of the archipelago. This study represents the first attempt to document COI barcodes for soil mesofauna in the Canary Islands and provides the basis for the Canary Islands Soil Biodiversity barcode database (CISoilBiota). The wider implementation of this barcoding workflow within the Canaries holds the promise for a massive biodiversity discovery.

Acknowledgements

Conflicts of interest

The authors have declared that no competing interests exist.

References

- André HM, Noti MI, Lebrun P (1994) The soil fauna: the other last biotic frontier. Biodiversity and Conservation 3 (1): 45-56. https://doi.org/10.1007/BF00115332
- Andújar C, Arribas P, Yu DW, Vogler AP, Emerson BC (2018) Why the COI barcode should be the community DNA metabarcode for the metazoa. Molecular Ecology 27 (20): 3968-397. https://doi.org/10.1111/mec.14844
- Andújar C, Arribas P, López H, Arjona Y, Pérez-Delgado A, Oromí P, Vogler AP, Emerson BC (2022) Community assembly and metaphylogeography of soil biodiversity: Insights from haplotype-level community DNAmetabarcoding within an oceanic island. Molecular Ecology 31 (15): 4078-4094. https://doi.org/10.1111/mec.16560
- Arribas P, Andújar C, Hopkins K, Shepherd M, Vogler AP (2016) Metabarcoding and mitochondrial metagenomics of endogean arthropods to unveil the mesofauna of the soil. Methods in Ecology and Evolution 7 (9): 1071-1081. https://doi.org/10.1111/2041-210X.12557
- Arribas P, Andújar C, Moraza ML, Linard B, Emerson BC, Vogler AP (2019) Mitochondrial metagenomics reveals the ancient origin and phylodiversity of soil mites and provides a phylogeny of the Acari. Molecular Biology and Evolution 37 (3): 683-694. https://doi.org/10.1093/molbev/msz255
- Arribas P, Andújar C, Bohmann K, DeWaard JR, Economo EP, Elbrecht V, Geisen S, Goberna M, Krehenwinkel H, Novotny V, Zinger L, Creedy TJ, Meramveliotakis E, Noguerales V, Overcast I, Morlon H, Papadopoulou A, Vogler AP, Emerson BC (2022) Toward global integration of biodiversity big data: a harmonized metabarcode data generation module for terrestrial arthropods. GigaScience 11: 1-12.
- Ashfaq M, Akhtar S, Rafi MA, Mansoor S, Hebert PD (2017) Mapping global biodiversity connections with DNA barcodes: Lepidoptera of Pakistan. PLOS One 12 (3): 1-13.
- Bardgett RD (2002) Causes and consequences of biological diversity in soil. Zoology 105 (4): 367-375. https://doi.org/10.1078/0944-2006-00072
- Bardgett RD, Van Der Putten WH (2014) Belowground biodiversity and ecosystem functioning. Nature 515 (7528): 505-511. https://doi.org/10.1038/nature13855
- Cameron EK, Martins IS, Lavelle P, Mathieu J, Tedersoo L, Gottschall F, Guerra CA, Hines J, Patoine G, Siebert J, Winter M, Cesarz S, Delgado-Baquerizo M, Ferlian O, Fierer N, Kreft H, Lovejoy TE, Montanarella L, Orgiazzi A, Pereira HM, Phillips HRP, Settele J, Wall DH, Eisenhauer N (2018) Global gaps in soil biodiversity data. Nature Ecology & Evolution 2 (7): 1042-104. https://doi.org/10.1038/s41559-018-0573-8
- Caterino M, Recuero E (2023) Molecular diversity of Protura in southern High Appalachian leaf litter. Biodiversity Data Journal 11.
- Chan A, Chiang LP, Hapuarachchi HC, Tan CH, Pang SC, Lee R, Lee KS, Ng LC, Lam-Phua SG (2014) DNA barcoding: complementing morphological identification of mosquito species in Singapore. Parasites & Vectors 7 (1): 1-12.
- Cicconardi F, Nardi F, Emerson BC, Frati F, Fanciulli PP (2010) Deep phylogeographic divisions and long-term persistence of forest invertebrates (Hexapoda: Collembola) in the North-Western Mediterranean basin. Molecular Ecology 19 (2): 386-400. https://doi.org/10.1111/j.1365-294X.2009.04457.x

- Cicconardi F, Fanciulli PP, Emerson BC (2013) Collembola, the biological species concept and the underestimation of global species richness. Molecular Ecology 22 (21): 5382-5396. https://doi.org/10.1111/mec.12472
- Cicconardi F, Borges PA, Strasberg D, Oromí P, López H, Pérez-Delgado AJ, Casquet J, Caujapé-Castells J, Fernández-Palacios JM, Thébaud C, Emerson BC (2017) MtDNA metagenomics reveals large-scale invasion of belowground arthropod communities by introduced species. Molecular Ecology 26 (12): 3104-3115. https://doi.org/10.1111/mec.14037
- Copilaş-Ciocianu D, Rewicz T, Sands AF, Palatov D, Marin I, Arbačiauskas K, Hebert P, Grabowski M, Audzijonyte A (2022) A DNA barcode reference library for endemic Ponto-Caspian amphipods. Scientific reports 12 (1).
- Creedy TJ, Andújar C, Meramveliotakis E, Noguerales V, Overcast I, Papadopoulou A, Morlon H, Vogler AP, Emerson BC, Arribas P (2021) Coming of age for COI metabarcoding of whole organism community DNA: Towards bioinformatic harmonisation. Molecular Ecology Resources 22 (3): 847-861. https://doi.org/10.1111/1755-0998.13502
- Decaëns T, Jiménez JJ, Gioia C, Measey GJ, Lavelle P (2006) The values of soil animals for conservation biology. European Journal of Soil Biology 42: S23-S38. https://doi.org/10.1016/j.ejsobi.2006.07.001
- Decaëns T (2010) Macroecological patterns in soil communities. Global Ecology and Biogeography 19 (3): 287-302. https://doi.org/10.1111/j.1466-8238.2009.00517.x
- deWaard JR, et al. (2019) A reference library for Canadian invertebrates with 1.5 million barcodes, voucher specimens, and DNA samples. Scientific Data 6 (1): 308. https://doi.org/10.1038/s41597-019-0320-2
- Emerson BC, Cicconardi F, Fanciulli PP, Shaw PJ (2011) Phylogeny, phylogeography, phylobetadiversity and the molecular analysis of biological communities. Philosophical Transactions of the Royal Society B: Biological Sciences 366 (1576): 2391-240. https://doi.org/10.1098/rstb.2011.0057
- Emerson BC, Borges PAV, Cardoso P, Convey P, DeWaard JR, Economo EP, Gillespie RG, Kennedy S, Krehenwinkel H, Meier R, Roderick GK, Strasberg D, Thébaud C, Traveset A, Creedy TJ, Meramveliotakis E, Noguerales V, Overcast I, Morlon H, Papadopoulou A, Vogler AP, Arribas P, Andújar C (2022) Collective and harmonized high throughput barcoding of insular arthropod biodiversity: Toward a genomic observatories network for islands. Molecular Ecology 00: 1-16.
- Endlweber K, Schädler M, Scheu S (2006) Effects of foliar and soil insecticide applications on the collembolan community of an early set-aside arable field. Applied Soil Ecology 31 (1-2): 136-146. https://doi.org/10.1016/j.apsoil.2005.03.004
- FAO, ITPS, GSBI, SCBD, EC (2020) State of knowledge of soil biodiversity - Status, challenges and potentialities. FAO, Rome, 618 pp. [ISBN 978-92-5-133582-6]
- Fattorini S (2009) On the general dynamic model of oceanic island biogeography. Journal of Biogeography 36 (6): 1100-1110. https://doi.org/10.1111/j.1365-2699.2009.02083.x
- 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 (5): 294-299.
- Gilyarov MS, Breguetova NG (1977) A key to the soilinhabiting mites, Mesostigmata. Akad. Nauk. Zool. Inst., Leningrad, Moscow, 537 pp. [In Russian].

- Gilyarov MS (1978) A key to the soil-inhabiting mites of the Trombidiformes. Akad. Nauk. Zool. Inst., Leningrad, Moscow, 260 pp. [In Russia].
- Guerra CA, Heintz-Buschart A, Sikorski J, Chatzinotas A, Guerrero-Ramírez N, Cesarz S, Beaumelle L, Rillig MC, Maestre FT, Delgado-Baquerizo M, Buscot F, Overmann J, Patoine G, Phillips HRP, Winter M, Wubet T, Küsel K, Bardgett RD, Cameron EK, Cowan D, Grenbenc T, Marín C, Orgiazzi A, Singh BK, Wall DH, Eisenhauer N (2020) Blind spots in global soil biodiversity and ecosystem function research. Nature Communications 11 (1): 3870. https://doi.org/10.1038/s41467-020-17688-2
- Hebert PD, Ratnasingham S, deWaard JR (2003a) Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. Proceedings of the Royal Society of London. Series B: Biological Sciences 270: S96-S99.
- Hebert PD, Cywinska A, Ball SL, DeWaard JR (2003b) Biological identifications through DNA barcodes. Proceedings of the Royal Society of London. Series B: Biological Sciences 270 (1512): 313-321. https://doi.org/10.1098/rspb.2002.2218
- Hendrich L, Morinière J, Haszprunar G, Hebert PD, Hausmann A, Köhler F, Balke M (2015) A comprehensive DNA barcode database for Central European beetles with a focus on Germany: adding more than 3500 identified species to BOLD. Molecular Ecology Resources 15 (4): 795-818. https://doi.org/10.1111/1755-0998.12354
- Jackson JK, Battle JM, White BP, Pilgrim EM, Stein ED, Miller PE, Sweeney BW (2014) Cryptic biodiversity in streams: a comparison of macroinvertebrate communities based on morphological and DNA barcode identifications. Freshwater Science 33 (1): 312-324. https://doi.org/10.1086/675225
- Janzen DH, Burns JM, Cong Q, Hallwachs W, Dapkey T, Manjunath R, Hajibabaei M, Hebert PDN, Grishin NV (2017) Nuclear genomes distinguish cryptic species suggested by their DNA barcodes and ecology. Proceedings of the National Academy of Sciences 114 (31): 8313-8318. https://doi.org/10.1073/pnas.1621504114
- Juan C, Emerson BC, Oromí P, Hewitt GM (2000) Colonization and diversification: towards a phylogeographic synthesis for the Canary Islands. Trends in Ecology & Evolution 15 (3): 104-109. https://doi.org/10.1016/S0169-5347(99)01776-0
- Kalyaanamoorthy S, Minh BQ, Wong TK, Von Haeseler A, Jermiin LS (2017) ModelFinder: fast model selection for accurate phylogenetic estimates. Nature Methods 14 (6): 587-589. https://doi.org/10.1038/nmeth.4285
- Karg W (1993) Acari (Acarina), Milben, Parasitiformes (Anactinochaeta), Cohors Gamasina Leach. Raubmilben. Die Tierwelt Deutschlands und der angrenzenden Meeresteile. Gustav Fischer Verlag, 523 pp.
- Katoh K, Misawa K, Kuma KI, Miyata T (2002) MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Research 30 (14): 3059-3066. https://doi.org/10.1093/nar/gkf436
- Kimura M (1979) The neutral theory of molecular evolution. Scientific American 241 (5): 98-129. https://doi.org/10.1038/scientificamerican1179-98
- Koh LP, Sodhi NS, Tan HTW, Peh KSH (2002) Factors affecting the distribution of vascular plants, springtails, butterflies and birds on small tropical islands. Journal of Biogeography 29 (1): 93-108. https://doi.org/10.1046/j.1365-2699.2002.00657.x
- Krantz GW, Walter DE (2009) A manual of acarology. 3rd Edition. Texas Tech University Press, Lubbock, Texas, 816 pp. [ISBN 9780896726208]
- Leo SS, Pybus MJ, Sperling FA (2010) Deep mitochondrial DNA lineage divergences within Alberta populations of *Dermacentor albipictus* (Acari: Ixodidae) do not indicate

distinct species. Journal of Medical Entomology 47 (4): 565-574. https://doi.org/10.1603/ME10006

- Magoga G, Fontaneto D, Montagna M (2021) Factors affecting the efficiency of molecular species delimitation in a species-rich insect family. Molecular Ecology Resources 21 (5): 1475-1489. https://doi.org/10.1111/1755-0998.13352
- Maraun M, Schatz H, Scheu S (2007) Awesome or ordinary? Global diversity patterns of oribatid mites. Ecography 30 (2): 209-216. https://doi.org/10.1111/j.0906-7590.2007.04994.x
- Marconi M, Modesti A, Alvarez LP, Ogoña PV, Mendoza AC, Vecco-Giove CD, Luna JO, Di Giulio A, Mancini E (2022) DNA barcoding of stingless bees (Hymenoptera: Meliponini) in northern Peruvian forests: A plea for integrative taxonomy. Diversity 14 (8): 632. https://doi.org/10.3390/d14080632
- Meier R, Blaimer BB, Buenaventura E, Hartop E, vonRintelen T, Srivathsan A, Yeo D (2022) A re-analysis of the data in Sharkey et al.'s (2021) minimalist revision reveals that BINs do not deserve names, but BOLD Systems needs a stronger commitment to open science. Cladistics 38 (2): 264-275. https://doi.org/10.1111/cla.12489
- Minh BQ, Nguyen MAT, Von Haeseler A (2013) Ultrafast approximation for phylogenetic bootstrap. Molecular Biology and Evolution 30 (5): 1188-1195. https://doi.org/10.1093/molbev/mst024
- Nielsen UN, Wall DH, Six J (2015) Soil biodiversity and the environment. Annual review of environment and resources 40: 63-90. https://doi.org/10.1146/annurev-environ-102014-021257
- Pérez-Delgado AJ, Arribas P, Hernando C, López H, Arjona Y, Suárez-Ramos D, Emerson BC, Andújar C (2022) Hidden island endemic species and their implications for cryptic speciation within soil arthropods. Journal of Biogeography 49 (7): 1367-1380. https://doi.org/10.1111/jbi.14388
- Pérez-Íñigo C (1993) Acari, Oribatei, Poronota. In: Ramos MA (Ed.) Fauna Ibérica. 3 Vol. Museo de Ciencias Naturales, CSIC, Madrid, 320 pp.
- Pérez-Íñigo C (1997) Acari, Oribatei, Gymnonota I. In: Ramos MA (Ed.) Fauna Ibérica. 9 Vol. Museo de Ciencias Naturales, CSIC, Madrid, 374 pp.
- Petersen H, Luxton M (1982) A comparative analysis of soil fauna populations and their role in decomposition processes. Oikos288-388. https://doi.org/10.2307/3544689
- Pfingstl T, Lienhard A, Baumann J, Koblmüller S (2021) A taxonomist's nightmare – Cryptic diversity in Caribbean intertidal arthropods (Arachnida, Acari, Oribatida). Molecular Phylogenetics and Evolution 163.
- Ponge JF (2013) Plant–soil feedbacks mediated by humus forms: a review. Soil Biology and Biochemistry 57: 1048-1060. https://doi.org/10.1016/j.soilbio.2012.07.019
- Porter TM, Hajibabaei M (2018) Over 2.5 million COI sequences in GenBank and growing. PLoS ONE 13 (9).
- Puillandre N, Brouillet S, Achaz G (2020) ASAP: assemble species by automatic partitioning. Molecular Ecology Resources 21 (2): 609-620. https://doi.org/10.1111/1755-0998.13281
- Ranasinghe UG, Eberle J, Thormann J, Bohacz C, Benjamin SP, Ahrens D (2022) Multiple species delimitation approaches with COI barcodes poorly fit each other and morphospecies–An integrative taxonomy case of Sri Lankan Sericini chafers (Coleoptera: Scarabaeidae). Ecology and Evolution 12 (5).

- Ratnasingham S, Hebert PD (2007) BOLD: The Barcode of Life Data System (http://www.barcodinglife.org). Molecular Ecology Notes 7 (3): 355-364. https://doi.org/10.1111/j.1471-8286.2007.01678.x
- Ronquist F, Forshage M, Häggqvist S, Karlsson D, Hovmöller R, Bergsten J, Holston K, Britton T, Abenius J, Andersson B, Neerup Buhl P, Coulianos C, Fjellberg A, Gertsson C, Hellqvist S, Jaschhof M, Kjærandsen J, Klopfstein S, Kobro S, Liston A, Meier R, Pollet M, Riedel M, Roháček J, Schuppenhauer M, Stigenberg J, Struwe I, Taeger A, Ulefors S, Varga O, Withers P, Gärdenfors U (2020) Completing Linnaeus's inventory of the Swedish insect fauna: Only 5,000 species left? PLoS ONE 15 (3).
- Salces-Castellano A, Andújar C, López H, Pérez-Delgado AJ, Arribas P, Emerson B (2021) Flightlessness in insects enhances diversification and determines assemblage structure across whole communities. Proceedings of the Royal Society B 288.
- Somervuo P, Yu DW, Xu CC, Ji Y, Hultman J, Wirta H, Ovaskainen O (2017) Quantifying uncertainty of taxonomic placement in DNA barcoding and metabarcoding. Methods in Ecology and Evolution 8 (4): 398-407. https://doi.org/10.1111/2041-210X.12721
- Srivathsan A, Baloğlu B, Wang W, Tan WX, Bertrand D, Ng AHQ, Boey EJH, Koh JJY, Nagarajan N, Meier R (2018) A MinION™-based pipeline for fast and cost-effective DNA barcoding. Molecular Ecology Resources, 18 (5): 1035-1049. https://doi.org/10.1111/1755-0998.12890
- Srivathsan A, Hartop E, Puniamoorthy J, Lee WT, Kutty SN, Kurina O, Meier R (2019) Rapid, large-scale species discovery in hyperdiverse taxa using 1D MinION sequencing. BMC Biology 17 (96): 1-20.
- Srivathsan A, Lee L, Katoh K, Hartop E, Kutty SN, Wong J, Yeo D, Meier R (2021) ONTbarcoder and MinION barcodes aid biodiversity discovery and identification by everyone, for everyone. BMC Biology 19: 1-21.
- Srivathsan A, Feng V, Suárez D, Meier R (2023) Rapid species discovery and identification with real-time barcoding facilitated by ONTbarcoder 2.0 and Oxford Nanopore R10.4. bioRxiv https://doi.org/10.1101/2023.06.26.546538
- Suárez D, Arribas P, Jiménez-García E, Emerson BC (2022) Dispersal ability and its consequences for population genetic differentiation and diversification. Proceedings of the Royal Society B 289 (1975).
- Suárez D, Arribas P, Macías-Hernández N, Emerson B (2023) Dispersal ability and niche breadth influence interspecific variation in spider abundance and occupancy. Royal Society Open Science 10 (5).
- Szudarek-Trepto N, Kazmierski A, Dabert J (2021) Long-term stasis in acariform mites provides evidence for morphologically stable evolution: Molecular vs. morphological differentiation in *Linopodes* (Acariformes; Prostigmata). Molecular Phylogenetics and Evolution 163.
- Trifinopoulos J, Nguyen LT, von Haeseler A, Minh B (2016) W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. Nucleic Acids Research 44 (W1): W232-W235. https://doi.org/10.1093/nar/gkw256
- Vasilita C, Feng V, Hansen AK, Hartop E, Srivathsan A, Meier R (2023) Express barcoding with NextGenPCR and MinION for species-level sorting of ecological samples. bioRxiv.
- Veresoglou SD, Halley JM, Rillig MC (2015) Extinction risk of soil biota. Nature Communications 6 (1).

- Walter DE, Proctor HC (2013) Mites: ecology, evolution and behaviour. Springer, 494 pp.
- White HJ, León-Sánchez L, Burton VJ, Cameron EK, Caruso T, Cunha L, Dirilgen T, Jurburg SD, Kelly R, Kumaresan D, Ochoa-Hueso R, Ordonez A, Phillips HRP, Prieto I, Schmidt O, Caplat P (2020) Methods and approaches to advance soil macroecology. Global Ecology and Biogeography 29 (10): 1674-1690. https://doi.org/10.1111/geb.13156
- Yin Y, Yao LF, Hu Y, Shao ZK, Hong XY, Hebert PD, Xue XF (2022) DNA barcoding uncovers cryptic diversity in minute herbivorous mites (Acari, Eriophyoidea). Molecular Ecology Resources 22 (5): 1986-199. https://doi.org/10.1111/1755-0998.13599
- Young MR, Behan-Pelletier VM, Hebert PD (2012) Revealing the hyperdiverse mite fauna of subarctic Canada through DNA barcoding. PLoS ONE 7 (11).
- Young MR, Proctor HC, deWaard JR, Hebert PD (2019) DNA barcodes expose unexpected diversity in Canadian mites. Molecular Ecology 28 (24).
- Yu DW, Ji Y, Emerson BC, 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
- Zhang J, Kapli P, Pavlidis P, Stamatakis AA (2013) A general species delimitation method with applications to phylogenetic placements. Bioinformatics 29 (22): 2869-2876. https://doi.org/10.1093/bioinformatics/btt499

Supplementary materials

Suppl. material 1: Table S1 doi

Authors: Irene Santos-Perdomo, Daniel Suárez, María L. Moraza, Paula Arribas, Carmelo Andújar
Data type: Dataset (excel table)
Brief description: Sampling localities, with data on habitat type, coordinates, altitude and date of collection.
Download file (10.64 kb)

Suppl. material 2: Figure S1 doi

Authors: Irene Santos-Perdomo, Daniel Suárez, María L. Moraza, Paula Arribas, Carmelo Andújar
Data type: Image (TIFF)
Brief description: Maximum Likelihood (ML) phylogenetic tree obtained for the barcode dataset with IQ-TREE using the best fitting substitution model for each codon partition and nodal support obtained by 1,000 ultrafast boot-strap replicates.
Download file (5.79 MB)

Suppl. material 3: Table S2 `doi`

Authors: Irene Santos-Perdomo, Daniel Suárez, María L. Moraza, Paula Arribas, Carmelo Andúja

Data type: Dataset (excel table)

Brief description: Specimens morphological identification at order, family, genus and species levels.

Download file (14.49 kb)

Suppl. material 4: Table S3 `doi`

Authors: Irene Santos-Perdomo, Daniel Suárez, María L. Moraza, Paula Arribas, Carmelo Andújar

Data type: Dataset (excel table)

Brief description: Data of studied specimens within the BOLD project CISoilBiota, subproject CIACA (Acari of the Canary Islands).

Download file (21.49 kb)

Suppl. material 5: Figure S2a `doi`

Authors: Irene Santos-Perdomo, Daniel Suárez, María L. Moraza, Paula Arribas, Carmelo Andújar

Data type: Phylogenetic tree in PDF format

Brief description: Distance-based phylogenetic tree generated by BOLD using K2P corrected distances including best BOLD matches of *Acrogalumna longipluma* (lineage A).

Download file (30.27 kb)

Suppl. material 6: Figure S2b `doi`

Authors: Irene Santos-Perdomo, Daniel Suárez, María L. Moraza, Paula Arribas, Carmelo Andújar

Data type: Phylogenetic tree in PDF forma

Brief description: Distance-based phylogenetic tree generated by BOLD using K2P corrected distances including best BOLD matches of *Acrogalumna longipluma* (lineage B).

Download file (30.88 kb)

Suppl. material 7: Figure S3 `doi`

Authors: Irene Santos-Perdomo, Daniel Suárez, María L. Moraza, Paula Arribas, Carmelo Andújar

Data type: Phylogenetic tree in PDF format

Brief description: Distance-based phylogenetic tree generated by BOLD using K2P corrected distances including best BOLD matches of *Hypochthonius luteus*.

Download file (30.88 kb)

Suppl. material 8: Figure S4 doi

Authors: Irene Santos-Perdomo, Daniel Suárez, María L. Moraza, Paula Arribas, Carmelo Andújar

Data type: Phylogenetic tree in PDF format

Brief description: Distance-based phylogenetic tree generated by BOLD using K2P corrected distances including best BOLD matches of *Odontocepheus elongatus*.

Download file (31.96 kb)

Suppl. material 9: Figure S5a doi

Authors: Irene Santos-Perdomo, Daniel Suárez, María L. Moraza, Paula Arribas, Carmelo Andújar

Data type: Phylogenetic tree in PDF format

Brief description: Distance-based phylogenetic tree generated by BOLD using K2P corrected distances including best BOLD matches of *Acrotritia ardua ardua*.

Download file (30.92 kb)

Suppl. material 10: Figure S5b doi

Authors: Irene Santos-Perdomo, Daniel Suárez, María L. Moraza, Paula Arribas, Carmelo Andújar

Data type: Phylogenetic tree in PDF format

Brief description: Distance-based phylogenetic tree generated by BOLD using K2P corrected distances including best BOLD matches of *Acrotritia penicillata*.

Download file (30.10 kb)

Suppl. material 11: Figure S6 doi

Authors: Irene Santos-Perdomo, Daniel Suárez, María L. Moraza, Paula Arribas, Carmelo Andújar

Data type: Phylogenetic tree in PDF format

Brief description: Distance-based phylogenetic tree generated by BOLD using K2P corrected distances including best BOLD matches of *Phthiracarus cf globosus*.

Download file (30.10 kb)

Suppl. material 12: Figure S7 doi

Authors: Irene Santos-Perdomo, Daniel Suárez, María L. Moraza, Paula Arribas, Carmelo Andújar

Data type: Phylogenetic tree in PDF format

Brief description: Distance-based phylogenetic tree generated by BOLD using K2P corrected distances including best BOLD matches of *Xenillus tegeocranus*.

Download file (30.07 kb)