



Environmental DNA of insects and springtails from caves reveals complex processes of eDNA transfer in soils

Enrico Lunghi ^{a,b,*}, Barbara Valle ^{c,d}, Alessia Guerrieri ^e, Aurélie Bonin ^e, Fabio Cianferoni ^{f,g}, Raoul Manenti ^{e,h}, Gentile Francesco Ficetola ^{e,i}

^a Division of Molecular Biology Ruder Bošković Institute, Zagreb, Croatia

^b Natural Oasis, Prato, Italy

^c Dipartimento di Bioscienze, Università degli Studi di Milano, Milano, Italy

^d Unità di Climatologia ed Ecologia, MUSE-Museo delle Scienze di Trento, Italy

^e Dipartimento di Scienze e Politiche Ambientali, Università degli Studi di Milano, Milano, Italy

^f Istituto di Ricerca sugli Ecosistemi Terrestri (IRET), Consiglio Nazionale delle Ricerche (CNR), Sesto Fiorentino (Firenze), Italy

^g Zoologia, La Specola, Museo di Storia Naturale, Università degli Studi di Firenze, Firenze, Italy

^h Laboratorio di Biologia Sotterranea "Enrico Pezzoli", Parco Regionale del Monte Barro, Galbiate, Italy

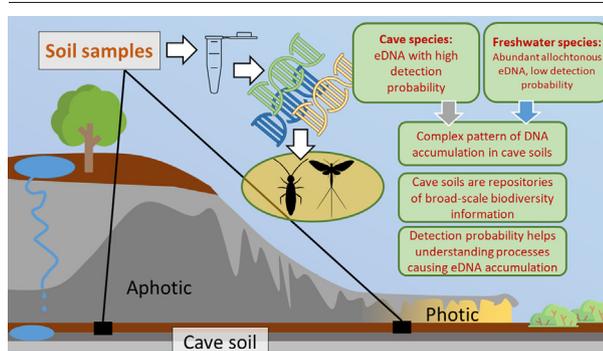
ⁱ Laboratoire d'Écologie Alpine (LECA), Université Grenoble Alpes, CNRS, Grenoble, France



HIGHLIGHTS

- We tested the usefulness of eDNA from cave soil to study subterranean biodiversity.
- A complex pattern of allochthonous and autochthonous eDNA has been observed.
- Cave soils act as conveyor belts of biodiversity information.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:

Received 13 December 2021

Received in revised form 1 February 2022

Accepted 15 February 2022

Available online 22 February 2022

Editor: Fang Wang

Keywords:

Biodiversity
Biospeleology
Cave biology
DNA metabarcoding
Monitoring methodology
Subterranean

ABSTRACT

Subterranean environments host a substantial amount of biodiversity, however assessing the distribution of species living underground is still extremely challenging. Environmental DNA (eDNA) metabarcoding is a powerful tool to estimate biodiversity in poorly known environments and has excellent performance for soil organisms. Here, we tested 1) whether eDNA metabarcoding from cave soils/sediments allows to successfully detect springtails (Hexapoda: Collembola) and insects (Hexapoda: Insecta); 2) whether eDNA mostly represents autochthonous (cave-dwelling) organisms or it also incorporates information from species living in surface environments; 3) whether eDNA detection probability changes across taxa with different ecology. Environmental DNA metabarcoding analyses detected a large number of Molecular Operational Taxonomic Units (MOTUs) for both insects and springtails. For springtails, detection probability was high, with a substantial proportion of hypogean species, suggesting that eDNA provides good information on the distribution of these organisms in caves. Conversely, for insects most of MOTUs represented taxa living outside caves, and the majority of them represented taxa/organisms living in freshwater environments (Ephemeroptera, Plecoptera and Trichoptera). The eDNA of freshwater insects was particularly abundant in deep sectors of caves, far from the entrance. Furthermore, average detection probability of insects was significantly lower than the one of springtails. This suggests that cave soils/sediments act as “conveyor belts of biodiversity information”, possibly

* Corresponding author at: Division of Molecular Biology, Ruder Bošković Institute, Bijenicka cesta 54, 10000 Zagreb, Croatia.
E-mail address: enrico.arti@gmail.com (E. Lunghi).

because percolating water lead to the accumulation of eDNA of organisms living in nearby areas. Cave soils hold a complex mix of autochthonous and allochthonous eDNA. eDNA provided unprecedented information on the understudied subterranean cave organisms; analyses of detection probability and occupancy can help teasing apart local eDNA from the eDNA representing spatially-integrated biodiversity for whole landscape.

1. Introduction

The subterranean domain hosts an impressive but underappreciated amount of biodiversity. It has been estimated that there are >500,000 km³ of underground spaces and cavities just in the United States, and that up to 50% of the animal biomass may live below the terrestrial surface (Culver and Pipan, 2009; Fierer et al., 2009). The subterranean domain is composed by a complex mosaic of elements. We usually refer to subterranean environments with the term “caves”, but caves are just the underground spaces that are large enough to potentially allow human access; besides these, a very broad range of environments, such as interstitial habitats or groundwaters, are inhabited by organisms but remain essentially inaccessible to humans (Culver and Pipan, 2009; Romero, 2009; Ficetola et al., 2019a; Mammola, 2019; Canedoli et al., 2022). As a consequence, the cave-dwelling organisms are among the least known components of Earth's biodiversity. Entering caves and exploring them is usually challenging (MacNeil and Brcic, 2017; Mammola et al., 2021; Canedoli et al., 2022), thus most of the subterranean domain remains unexplored, and we cannot describe, analyze or conserve the biodiversity of the environments we never explored (Ficetola et al., 2019a). Adopting standardized monitoring protocols for a quantitative assessment of biodiversity can be even more challenging, given the logistical constraints associated to these environments (Culver et al., 2012; Mammola et al., 2021). Therefore, we urgently need new approaches to better understand the functioning of subterranean ecosystems and to fully exploit their potential in addressing evolutionary and ecological questions (Mammola, 2019; Mammola et al., 2020; Mammola et al., 2021; Canedoli et al., 2022).

Environmental DNA (eDNA) metabarcoding has been proposed as a new tool to unveil the biology of underground fauna (Ficetola et al., 2019b; Saccò et al., 2019). eDNA extracted from groundwater has been successfully used to study multiple organisms, including fishes, amphibians and crayfishes, and allowed detecting target species in new, unexplored sites (Goricki et al., 2017; Vörös et al., 2017; Saccò et al., 2019; Boyd et al., 2020; West et al., 2020; White et al., 2020). So far, eDNA studies in underground environments mostly focused on freshwater organisms and extracted eDNA from water (Vörös et al., 2017; Saccò et al., 2019; Boyd et al., 2020; West et al., 2020; White et al., 2020). However, eDNA can be successfully extracted also from soils and sediments, and soil eDNA has allowed the characterization of communities of animals living both below (e.g. earthworms, springtails, nematodes; Bienert et al., 2012; Zinger et al., 2019b; Rota et al., 2020; Rosero et al., 2021) and above soil surface (e.g. mammals, birds, insects; Andersen et al., 2012; Slon et al., 2017; Zinger et al., 2019b; Rosero et al., 2021).

Caves host an impressive number of terrestrial invertebrates, with many endemic and poorly studied species. eDNA collected from the cave soils and sediments (for the sake of brevity, hereafter cave soil) could be a great approach to assess these understudied organisms. However, the processes determining the transfer and deposition of eDNA from organisms to soil and sediments can be complex. In fact, eDNA can be transported by multiple processes, such as the movement of free water or sediments (Haile et al., 2007; Giguët-Covex et al., 2019). Caves are often carved by complex erosive processes, and the occurrence of non-autochthonous eDNA can make it challenging to study present or past communities associated to underground environments (Haile et al., 2007). Insects (Hexapoda: Insecta) and springtails (Hexapoda: Collembola) are among the most widespread animals on Earth, and have very broad ecological spectra. Springtails are small, wingless soil arthropods living in almost all terrestrial habitats; they are particularly abundant in soils, where they can attain high densities (Hopkin, 1997; Balian et al., 2007; Gibert and Culver, 2009). Springtails

have a major functional role in soil as they affect organic matter decomposition and feed on fungi, dead organic matter, algae and soil microorganisms (Maafß et al., 2015; Potapov et al., 2016). They have a broad range of specialization to underground life and include both species with morphological adaptations to deep underground environments (e.g. reduction/loss of eyes and pigmentation; hereafter, hypogean species), and species associated to surface environments with eyes and pigmentation (hereafter, epigeic species) (Gibert and Deharveng, 2002; Kováč et al., 2016; Howarth and Moldovan, 2018). Springtails are one of the taxa of Hexapoda with the highest percentage of troglobites (6.15% of described species) and they usually dominate terrestrial oligotrophic caves (Deharveng and Bedos, 2018). Insects hold the highest taxonomic richness on Earth, inhabit nearly all the non-marine environments, and include species living in outdoor terrestrial environments but also cave specialists, with many species spending a substantial part of their life cycle in freshwaters (Balian et al., 2007; Romero, 2009). Many insects living in outdoor environments can enter inside caves actively or passively (accidental species), constituting a key component of underground food webs and being important subsidies for the functioning of these ecosystems (Schneider et al., 2011; Lunghi et al., 2014). By merging ecological information on these taxonomic groups with data on their occurrence in cave soils we can explore the potential and appropriateness of eDNA for studying the biodiversity of these environments.

Here we used a metabarcoding approach to study soil eDNA of insects and springtails from multiple cave soils. First, we evaluated the proportion of freshwater, outdoor and cave specialists among taxa detected through eDNA metabarcoding. If eDNA mostly represents local communities, we predict that cave-dwelling species will constitute the majority of detected taxa and/or of eDNA reads. Conversely, if eDNA results from complex processes of transport, we expect an overrepresentation of outdoor and/or freshwater taxa. Second, we analyzed eDNA detection probability (Ficetola et al., 2015; Furlan et al., 2016; Lahoz-Monfort et al., 2016; Guillera-Aroita et al., 2017; Chen and Ficetola, 2019) to better understand the processes determining the occurrence of different taxonomic and ecological groups in cave eDNA. Taxa with higher local abundance tend to have a larger frequency of positive PCR replicates in samples where they are actually present, compared to sporadic species (Ficetola et al., 2016; Furlan et al., 2016; Lahoz-Monfort et al., 2016). Thus, we expect that taxa actually present in cave soil had a more consistent eDNA signal (e.g. they are detected in most PCR replicates of the same sample) and a higher detection probability compared to taxa whom eDNA only comes through leaching (Haile et al., 2007; Chen and Ficetola, 2020). Therefore, we predict that taxa representing allochthonous eDNA have lower detection probability, compared to taxa representing autochthonous eDNA.

2. Methods

2.1. Sampling

Between May and July 2016, we collected soil eDNA from the soil of 16 underground environments (natural or man-modified; hereafter, caves) in Italy (13 sites in Sardinia and three in mainland Italy; Table S1). Usually, none of the study caves is flooded, even though karstic processes have carved most of them. For each site, we collected two soil samples following the protocol described in Guerrieri et al. (2021). We divided each subterranean environment into two parts: the photic area (i.e. where light was present) and the aphotic area (i.e., complete darkness, where illuminance <0.01 lx). We established two one-m² plots, one six meters after the cave entrance (falling within the photic area) and one six meters after the beginning of the

aphotic area. In two cases, the photic area did not exist (a door applied to the entrance of these sites interrupts the incoming light), so we established one extra plot in the aphotic area. Furthermore, in two cases it was impossible to sample the aphotic area because the cave floor was made of gravel (supplementary Table S1). Within each plot, we collected four soil subsamples (depth: 0–20 cm) to form a pooled sample. Litter and coarse organic matter were excluded. All sampling equipment underwent strict decontamination protocols before the collection of each sample (treated at >1000 °C with a portable blow torch). Pooled samples were homogenized and from each of them, we took a subsample of 15 g of soil, which was immediately inserted in hermetic, sterile boxes with 40 g of silica gel. This approach allows a quick desiccation of soil and long-term storage of eDNA for metabarcoding analyses (Guerrieri et al., 2021).

In subterranean environments, soils are composed of both chemical and clastic sediments that are produced in loco or that are brought from external environments (White, 2007). Clastic sediments are usually brought and accumulated through water that flows into the subterranean space. These include old sediments, as their accumulation starts together with the formation of the subterranean passage, and usually ends when the water stops its subterranean flowing (Culver and Pipan, 2009). Depending on the type of connection with the external environment, the flowing or seeping of water from the surface can keep contributing in accumulating clastic sediments in the surrounding of the cave entrance or in specific passages (White, 2007; Popovic et al., 2020). Clastic sediments also comprise a significant amount of organic material, which can be locally produced (e.g. guano or dead organisms), brought by atmospheric agents and gravity, but also by the animals that actively move between the surface and subterranean environment (Schneider et al., 2011; Venarsky et al., 2014; Barzaghi et al., 2017; Lunghi et al., 2018). Organic sediments are generally comprised in the most superficial soil layer, as they are continuously produced and accumulated by the activity of cave-dwelling and cave-entering organisms (Salgado et al., 2014).

2.2. Molecular analyses and bioinformatics treatments

Environmental DNA from 30 samples was extracted using the NucleoSpin Soil Mini Kit (Macherey-Nagel) in a room dedicated to the analysis of soil eDNA, with an additional step where soil samples were mixed for 15 min with 20 ml of phosphate buffer (Taberlet et al., 2012). We also included four extraction controls (Zinger et al., 2019a). Springtail eDNA was amplified using the primers Coll01, which amplify a ~130 bp fragment of the 16S mitochondrial rDNA (Janssen et al., 2018). This marker amplifies springtail DNA with very few mismatches in the priming regions (100% of tested taxa amplified in silico by Taberlet et al., 2018); available data suggest it has an excellent taxonomic resolution (resolution at the genus level: 87%; Taberlet et al., 2018). Insect eDNA was amplified using the primers Inse01, which amplify a ~155 bp fragment of the 16S mitochondrial rDNA (Taberlet et al., 2018). These primers are similar to the ones designed by Clarke et al. (2014); *in-silico* and *in-vitro* analyses suggest they are able to successfully amplify >80% of insects and, if reference sequences are available, they often enable genus-level identification of insects (Ficetola et al., 2021). Primers included 8-nucleotide-long tags at the 5' end, each of which had at least five nucleotide differences with the other tags to enable discrimination of PCR replicates after sequencing (Taberlet et al., 2018). For each marker, we performed 45 PCR cycles with an annealing temperature of 51 °C (Coll01) or 52 °C (Inse01). Along with DNA extracts and extraction controls, we also included three amplification controls (containing the PCR mix but no DNA extract) and two blanks (no PCR mix, no DNA) (Zinger et al., 2019a). Amplification was performed using 2 µl of DNA extracts in a 20-µl volume with 2 µl of primers mix (5 µM of each primer), 10 µl of AmpliTaq Gold 360 Master Mix 2× (Applied Biosystems), and 3.2 µg of bovine serum albumin (Roche Diagnostics). All samples and controls underwent four PCR replicates (Ficetola et al., 2015). Amplicons from the same marker were pooled together, purified using the MinElute PCR Purification Kit (Qiagen) and sequenced using a Illumina HiSeq 2500 platform at Fasteris (Geneva, Switzerland; <https://www.fasteris.com/dna/>), as described in Ficetola et al. (2021).

Sequence data were treated using the OBITools software (Boyer et al., 2016). Sequences were aligned using the *illumina-paired-end* program. Sequences with poorly aligned paired-ends (i.e. alignments with score < 40) were filtered out. Retained sequences were then assigned to samples (*ngsfilter* program) and dereplicated using the *obitool* program. Subsequently, using the *obiclean* program, we kept sequences that were at least twice as abundant as other related sequences differing by one nucleotide base in the same PCR replicate (“head sequences”), to remove PCR and sequencing errors. Retained sequences were then clustered at a 96% (Inse01) or 85% (Coll01) similarity threshold using the *sumacust* program (<https://git.metabarcoding.org/obitools/sumacust/wikis/home>) (Mercier et al., 2013). Preliminary bioinformatic analyses showed that these similarity thresholds minimize the risk that sequences attributed to the same species are clustered in different Molecular Operational Taxonomic Units (MOTUs), while limiting the over-merging of sequences belonging to distinct species (Bonin et al., 2021). The taxonomic assignment was performed on the basis of the EMBL reference database (version 140) using the *ecotag* program, following the procedure described in Guerrieri et al. (2021). We only retained MOTUs assigned to the target classes (Collembola or Insecta). Blanks, extraction and PCR controls were used to identify potential contaminants (Zinger et al., 2019a). On the basis of blanks, we only considered MOTUs detected with at least five reads per PCR replicate. We discarded as possible contaminants all MOTUs detected in more than one control with at least two replicates each, or in one control with >2 replicates. Overall, we discarded two MOTUs for springtails, and two MOTUs for insects as contaminants (0.8% and 0.3% of MOTUs discarded, respectively).

2.3. Ecological information on retained MOTUs

We used the available literature to extract ecological information on the obtained MOTUs of springtails and insects. Ecological information was extracted for all the MOTUs identified at the family level or at finer taxonomic groups. For springtails, MOTUs were assigned to either epigeal (pigmented organisms, with more than four eyes) or to hypogean (unpigmented organism, with four eyes or less, living below the soil surface) groups on the basis of available information on taxonomy, life history and ecology (see Supplementary Material for references). If a MOTU was identified at the genus or family level, the taxon received a score corresponding to the proportion of hypogean species with respect to the total number of species within that taxon on the basis of the last complete checklist of the Italian Fauna (Minelli and Stoch, 2006).

For insects, we considered the use of two non-exclusive habitat categories: freshwater (Y/N), and caves (Y/N). Freshwater species were defined as taxa who spend at least part of their lifecycle in freshwater: i.e. with aquatic larvae (e.g. Ephemeroptera, Plecoptera, Trichoptera) or both larvae and adults living in water (e.g. Coleoptera Elmidae). For insects, information on the actual frequency of a single species in underground environments is often lacking, partly because of the complexity of morphological identification for many of them (e.g. Lunghi et al., 2020). Cave MOTUs were thus defined as the taxa that are known to regularly exploit underground environments. A few insect MOTUs (0.9%) were unidentified families belonging to the Plecoptera and Ephemeroptera order. In the study area all the Plecoptera and Ephemeroptera taxa have aquatic larvae and none of them is a cave specialist, thus these MOTUs were grouped into the aquatic, non-cave specialist category. Taxonomy followed the NCBI reference taxonomic database, version 140 (<https://www.ncbi.nlm.nih.gov/taxonomy>). Information on the distribution and ecology of insect MOTUs was obtained from available checklists and on monographs (see Supplementary Material).

2.4. Statistical analyses

Debate is ongoing on the relevance of the frequency of reads assigned to a taxon as an estimator of its relative abundance. On the one hand, multiple factors can bias the relationship between frequency of reads and relative abundance (e.g. differences in match with primers, differences in eDNA shedding rate, GC content) (Fonseca, 2018; Nichols et al., 2018). Nevertheless, the

frequency of reads is often positively correlated to the relative abundance (e.g. biomass) of species (Yoccoz et al., 2012; Pansu et al., 2015; Evans et al., 2016), and recent analyses highlight that indexes weighting MOTUs on the basis of their relative abundance provide more robust estimates of actual biodiversity (Calderón-Sanou et al., 2020; Mächler et al., 2021), thus suggesting that the frequency of reads has a strong ecological meaning. As a consequence, in our analyses, we used approaches based on both the relative abundance and on the detection/non-detection of MOTUs.

First, we used linear mixed models (LMMs) to assess the relationships between MOTU frequency, habitat (hypogean vs epigean) and area of the cave. For each class (insects and springtails), we assessed whether the frequency of MOTUs from each habitat was different between the photic and the aphotic area of the cave. Cave area (photic / aphotic) was added as a fixed factor, while cave identity was included as a random factor. MOTU frequency was logit transformed after applying the [Smithson and Verkuilen \(2006\)](#) transformation to remove fixed 0's and 1's. Mixed models were run using the `lme4` and `lmerTest` packages in R 3.6.2; degrees of freedom were approximated following the Satterthwaite's method, so they can be not integer ([Bates et al., 2015](#); [Kuznetsova et al., 2017](#)).

Second, we used site occupancy-detection models (SODMs) to assess the detection probability of eDNA from different taxonomic and ecological groups. SODMs model the presence/absence (occupancy) of each MOTU in each sample, considering the possibility that present species are missed (false absences) and also considering the possibility of false detections. SODMs thus provide key information on the processes determining the detection / non-detection of eDNA ([Ficetola et al., 2015](#); [Lahoz-Monfort et al., 2016](#); [Guillera-Aroita et al., 2017](#); [Chen and Ficetola, 2019](#); [Chen and Ficetola, 2020](#)). We fitted a separate SODM for each MOTU detected in >3 independent PCR replicates; SODMs were run in a Bayesian framework assuming the frequency of false positives ([Lahoz-Monfort et al., 2016](#)). As prior for the frequency of false positives, we used a uniform distribution with interval (0– F_{max}). For each MOTU, F_{max} was the upper 90% CI of the binomial intervals of the frequency of that MOTU across all controls. With this approach, the frequency of false positives was informed by the frequency of contaminants detected in controls. We run three MCMC chains, 15,000 iterations per chain and a burn-in of 7500. These parameters ensured convergence for all the species (R-hat always ≤ 1.1). SODMs were run in jags using the `R2jags` package ([Su and Yajima, 2015](#)). The outcome of SODM can provide key information on concentration/dispersion of eDNA across samples ([Furlan et al., 2016](#); see Discussion). We thus used LMMs to assess whether the detection probability (i.e. the probability that eDNA is detected, if present in a sample) and the occupancy (i.e. the frequency of samples where eDNA is present) of eDNA is different between classes (springtails vs. insects) and, within each class, among MOTUs with different habitats. To take phylogenetic structure into account, family and order (i.e. taxonomy) were included in LMMs as random effects. Degrees of freedom were Satterthwaite approximated.

3. Results

After bioinformatics filtering, we obtained 2,510,000 reads for springtails and 820,000 reads for insects. Overall, we retained 156 MOTUs for springtails, and 439 MOTUs for insects. For springtails, we identified MOTUs from all four orders and from 10 different families ([Table 1](#)); 46 MOTUs were identifiable at the family level or better. For insects, we identified MOTUs from 11 orders and 44 different families; 263 MOTUs were identifiable at the family level or better. Remarkably, 42% of MOTUs were identified as Ephemeroptera (mayflies), Plecoptera (stoneflies) and Trichoptera (caddisflies), which have freshwater larvae and do not include cave-dwelling taxa in the study area ([Table 1](#)). Cave-dwelling insects only included cave crickets (Orthoptera: Rhaphidophoridae), and limonid crane flies (Diptera: Limoniidae). Limonids were the only group whose larvae are associated with freshwaters (aquatic), but whose adults often exploit caves ([Lunghi et al., 2020](#)). None of the detected springtails was a freshwater specialist.

For insects, the majority of reads belonged to species that exploit freshwater environments in at least one life-history stage (mean

Table 1

List of the MOTUs identified in this study. For springtails are indicated the percentage of hypogean MOTUs within the family, while for insects is indicated whether the MOTU is a cave dweller and whether belongs to freshwater environment.

Springtails Order	Family	N MOTUs	% hypogean		
Symphypleona	Bourletiellidae	1		0	
	Katiannidae	1		0	
Neelipleona	Neelidae	18		100	
Entomobryomorpha	Cyphoderidae	3		100	
	Entomobryidae	7		16	
	Isotomidae	7		12	
	Tomoceridae	3		14	
	Undefined family	15			
Poduromorpha	Hypogastruridae	3		8	
	Neanuridae	1		15	
	Onychiuridae	2		99	
Undefined order		95			
Insects			Cave species	Freshwater	
Odonata	Coenagrionidae	1	N	Y	
Orthoptera	Rhaphidophoridae	26	Y	N	
Coleoptera	Curculionidae	1	N	N	
	Staphylinidae	2	N	N	
	Elmidae	1	N	Y	
	Undefined family	1	-	-	
	Diptera	Sciaridae	3	N	N
	Psychodidae	1	Y	N*	
	Mycetophilidae	2	N	N	
	Cecidomyiidae	2	N	N	
	Tipulidae	1	Y	N	
	Limoniidae	1	Y	Y	
	Undefined family	1	-	-	
Hymenoptera	Braconidae	2	N	N	
	Apidae	2	N	N	
Hemiptera	Reduviidae	1	N	N	
	Pseudococcidae	12	N	N	
	Naucoridae	1	N	Y	
Ephemeroptera	Chizopteridae**	9	N	N	
	Baetidae	29	N	Y	
	Caenidae	3	N	Y	
	Ephemerellidae	5	N	Y	
	Ephemeridae	3	N	Y	
	Heptageniidae	23	N	Y	
	Isonychiidae**	1	N	Y	
	Leptophlebiidae	15	N	Y	
	Oligoneuriidae	1	N	Y	
	Potamanthidae	1	N	Y	
	Siphonuridae	2	N	Y	
	Undefined family	1	N	Y	
Phthiraptera	Menoponidae	1	N	N	
Plecoptera	Capniidae	1	N	Y	
	Chloroperlidae	5	N	Y	
	Leuctridae	16	N	Y	
	Nemouridae	14	N	Y	
	Perlidae	1	N	Y	
	Perlodidae	6	N	Y	
	Taeniopterygidae	4	N	Y	
	Undefined family	3	N	Y	
	Psocoptera	Epipsocidae	1	N	N
		Liposcelididae	1	N	N
		Psoquillidae**	11	N	N
		Undefined family	1	N	N
	Trichoptera	Leptoceridae	1	N	Y
		Limnephilidae	6	N	Y
Phryganeidae		2	N	Y	
Rhyacophilidae		12	N	Y	
Sericostomatidae		1	N	Y	
Undefined order		198	-	-	

* Some species of Psychodidae live in freshwaters though most of them in polluted water, which are absent close to the study sites.

** This family does not occur in Europe. The MOTU comes from a taxonomic group lacking in the reference database, probably related to this family.

frequency \pm SD: 93% \pm 13%), while just 2.5% \pm 2.4% belonged to cave-dwelling species ([Fig. 1](#)). The frequency of eDNA reads from freshwater insects was particularly high in the aphotic areas of caves

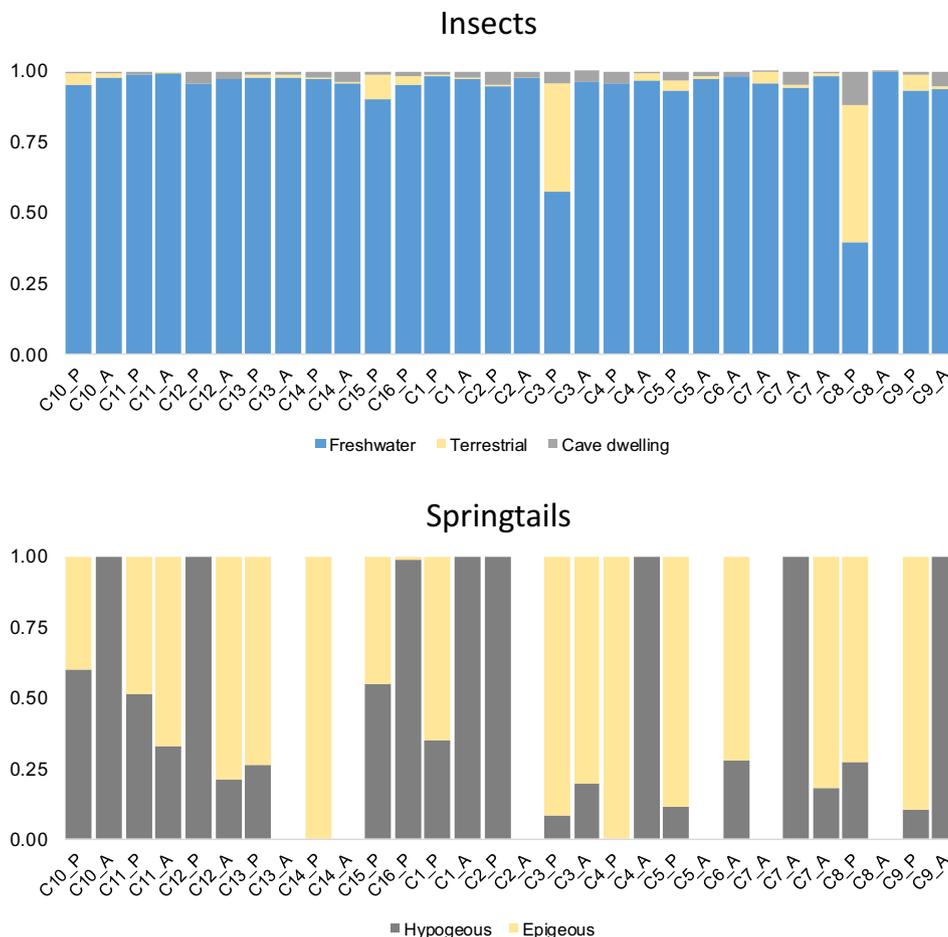


Fig. 1. Frequency of insect and springtail MOTUs with different ecological specialization across 30 soil samples from the cave floor. After cave codes, A indicates samples collected from the aphotic sector, while P indicates samples collected from the photic sector. Limoniidae are the only cave insects with freshwater larvae, and they have only been recorded from four samples; for easier plotting, we coded them as cave-dwelling.

(LMM: $F_{1,28} = 5.36$, $P = 0.028$; Fig. 2a), while we did not detect differences between the aphotic and the photic areas for cave-dwelling insects ($F_{1,22} = 1.77$, $P = 0.194$). For springtails, about half of

eDNA reads belonged to epigeal taxa and half belonged to hypogean taxa (mean frequency of each group: $50\% \pm 39\%$; Fig. 1). The frequency of reads from hypogean springtails was similar between the

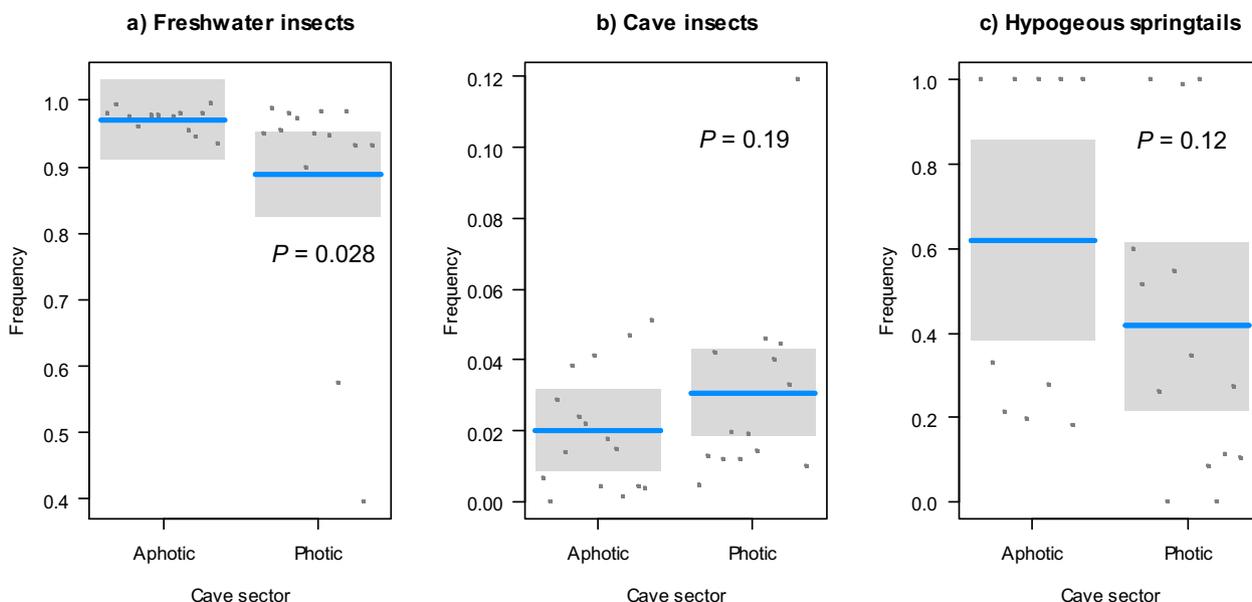


Fig. 2. Frequency of insects and springtails with different ecological specializations across cave sectors. Error bands are 95% confidence intervals.

aphotic and the photic areas ($F_{1,22} = 2.60$, $P = 0.123$). Results of LMM are summarized in Table S2.

3.1. Occupancy analyses

We run Bayesian occupancy models for the 239 insect MOTUs and 93 springtail MOTUs detected in >3 independent PCR replicates. Springtails showed significantly higher detection probability compared to insects (average detection probability: 0.67 ± 0.30 and 0.34 ± 0.32 for springtails and insects, respectively; mixed models taking into account taxonomic group: $F_{1,11.8} = 24.1$, $P < 0.001$; Fig. 3a). Conversely, the average occupancy of springtails was much lower than the one of insects (average occupancy: 0.13 ± 0.11 and 0.24 ± 0.14 for springtails and insects, respectively; $F_{1,11.3} = 13.4$, $P = 0.004$; Fig. 3c).

Within insects, freshwater MOTUs showed lower detection probability ($F_{1,6.96} = 9.2$, $P = 0.018$; Fig. 3b) and higher occupancy ($F_{1,153} = 30.5$, $P < 0.001$; Fig. 3d) compared to MOTUs from non-freshwater taxa, while we did not detect differences between cave-dwelling and surface-dwelling MOTUs for either detection probability and occupancy ($F_{1,11} = 3.0$, $P = 0.109$, and $F_{1,153} = 2.8$, $P = 0.09$). Within springtails, both average detection probability and occupancy were similar between epigeal and hypogean taxa ($F_{1,14.49} = 3.2$, $P = 0.095$ and $F_{1,4.5} = 2.1$, $P = 0.209$). Results of all LMM are summarized in Table S2.

4. Discussion

Our analyses revealed an unexpected pattern in eDNA representation within cave soils. The strong differences between insects and springtails, with a better detection probability of springtails, and the overrepresentation

of insects that show aquatic life stage, highlights the complexity of processes determining the transfer and accumulation of eDNA from organisms to soil.

4.1. Cave soils as “repositories of broad-scale biodiversity information”?

For insects, an unexpectedly high proportion of reads and MOTUs comes from freshwater specialists such as mayflies and stoneflies, that normally do not live in caves (Fig. 1). This strongly suggests that a substantial amount of insect eDNA is allochthonous, and originates from outdoor environments (White, 2007; Salgado et al., 2014). Although none of the study caves is regularly flooded (i.e. they are now inactive), water still percolates across these environments. It is thus likely that eDNA of freshwater insects originated from nearby streams and entered into the study caves through the intricate paths of water percolation that connects the two environments (White, 2007; Popovic et al., 2020). The complex movements of eDNA within and across environments remains a main open topic that needs to be addressed for a correct application of this approach to biodiversity surveys. Some studies suggested that eDNA from sediments and soils mostly has a local origin, and that it represents communities at the sampling point or just a few meters apart (Haile et al., 2007; Alsos et al., 2018; Zinger et al., 2019b). However, some analyses performed on surface waters observed that eDNA can be transported over broader distances, and proposed the idea that watercourses are “conveyor belts of biodiversity information”, as they can transfer eDNA across the drainage (Deiner et al., 2016). For instance, in Alpine systems, eDNA extracted from lake sediments represents the organisms living in the whole lake catchment, as the erosive processes can determine eDNA transfer from surface soils to rivers and then into the lake (Gigu et-Covex et al., 2019; Capo et al., 2021). Similarly, cave soils can contain the eDNA of organisms living both inside the cave and in nearby areas, and their analysis provided amazing environmental

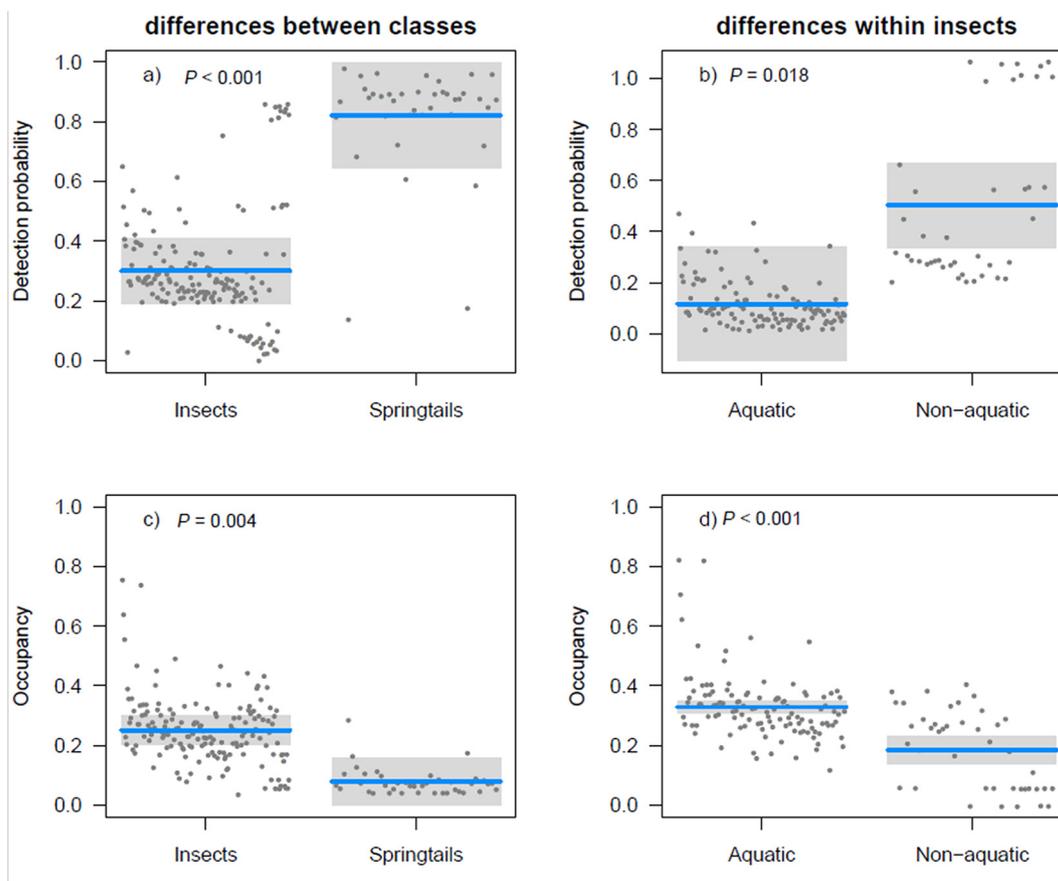


Fig. 3. Differences in detection probability between the eDNA of a) springtails and insects; b) aquatic and non-aquatic insects; and differences in occupancy between the eDNA of c) springtails and insects; c) aquatic and non-aquatic insects. Error bands are 95% confidence intervals.

reconstructions (Slon et al., 2017; Zhang et al., 2020). We took samples from the soil surface, therefore we do not expect that we sampled eDNA stored for long periods, even though more accurate stratigraphical information would be required to confirm the recent age of our samples (White, 2007).

Our study highlights the complex linkages between cave soils and surface waters. Strikingly, the frequency of eDNA from freshwater insects was particularly high in the aphotic zone of caves, i.e. far from cave entrance (Fig. 2a). In these sectors, the amount of accidental taxa incoming from the entrance of caves is limited (Lunghi et al., 2014; Lunghi et al., 2017), thus we expect less eDNA from accidental species, and this probably leads to a higher proportional representation of taxa for which eDNA is transported over long distances, for instance by water. Guano produced by bats is an additional potential source of DNA of surface insects (Salgado et al., 2014), even though none of the surveyed sites hosts a large bat colony. Conversely, it is unlikely that the observed pattern arose because of the primers selected, as the marker selected for the amplification of insects has limited amplification bias for most of orders, and is well represented in online databases (Ficetola et al., 2021). So far, the hypothesis of “conveyers belts of biodiversity information” has been mostly proposed for surface river networks (Deiner et al., 2016). Our study highlights that interconnections across ecosystems can be even more pervasive, as eDNA can move in multiple directions, from terrestrial to freshwater ecosystems and *viceversa*. Understanding the taphonomic processes affecting the origin of eDNA remains an open topic of molecular ecology research (Parducci et al., 2017; Capo et al., 2021).

4.2. Processes underlying eDNA detection in caves

The analysis of detection probability and occupancy, often performed through site occupancy-detection models (SODMs), is increasingly important to fully exploit biodiversity data including eDNA (Guillera-Arroita, 2017), and can help in understanding the processes underlying eDNA detection (Furlan et al., 2016; Martins et al., 2021). High values of detection probability indicate taxa that are consistently detected in most replicated analyses of the same samples, while taxa with low detection probability are only infrequently detected in the samples containing their eDNA. Furlan et al. (2016) proposed a framework to infer information on the concentration of eDNA on the basis of occupancy models. Taxa with clumped eDNA, reaching high concentration in some locations and not others, are expected to show high detection probability and low occupancy values. Conversely, taxa whose eDNA is randomly dispersed through the environment are expected to show low detection probability and high occupancy.

Our results match very well the Furlan et al. (2016) expectations. For instance, *Tomocerus* sp. (epigeal springtail) is a typical taxonomic unit with high detection probability ($P = 0.866$) and low occupancy, as it was detected in all the four PCR replicates of the single sample where it was found (Supplementary Table 1). Conversely, the insect MOTU identified as *Baetis* sp. (Ephemeroptera; MOTU Inse01_00170 in Supplementary Table 2) was detected in ten soil samples but, within each of them, it was only detected in one to two PCR replicates. As a consequence, *Baetis* sp. showed an extremely low detection probability ($P = 0.13$) and high occupancy ($\psi = 0.30$). This indicates sporadic detection of eDNA molecules randomly dispersed in the environment, which frequently arises when eDNA extracts only contain a few molecules of the target taxon. *Baetis* is an example of widespread mayfly genus, being among the commonest benthic insects in European streams and rivers (Balian et al., 2007; Bauernfeind and Soldan, 2012). The sporadic detection of these mayflies across multiple caves suggests that their eDNA can be present at very low concentration and high dispersion in percolating water and then in the cave soil (Furlan et al., 2016; Popovic et al., 2020). Within insects, detection probabilities are significantly higher for non-aquatic taxa (Fig. 3), and this suggests a more local origin, for instance from epigeal species living in the surrounding of the caves that sporadically use areas nearby the cave entrance as shelter (Fig. 2a; Lunghi et al., 2014; Lunghi et al., 2017). In the last years, it is increasingly evident that underground environments are important not

just for obligate cave dwellers, but also for a surprising number of surface species that regularly exploit them in specific seasons and provide a major subsidy of allochthonous biomass and nutrients to these resource-depleted environments (Schneider et al., 2011; Fiser et al., 2014).

The pattern was different for springtails, which showed consistently high detection probability, with similar patterns across taxa with different ecological traits, from hypogean to epigeal forms. This suggests a clumped eDNA distribution and a much more local signal compared to what is observed for both freshwater and ground-dwelling insects (Furlan et al., 2016). Thus, springtail presence in eDNA could not be related only to watercourses transport of eDNA from surface soils, but mostly represents autochthonous living organisms. Hypogean springtails are able to live in cave environment, and, when springtails do not show troglomorphy (Deharveng and Bedos, 2018), it is difficult to distinguish obligate cave dwellers from surface soil species (Dallai and Malatesta, 1982). In addition, many epigeal springtails colonize and exploit caves and other underground environments, where they can maintain stable populations (Kováč et al., 2016; Rendoš et al., 2016). In cave soils, the frequency of hypogean and epigeal species was similar, both in the aphotic and photic zones. The high detection probability of both groups suggests that the local eDNA signal is strong, and that ecological plasticity allows the two groups to colonize both surface soils and underground environments.

In our analyses, we did not focus on obligate cave-dwelling springtails, because the taxonomic resolution of our data was too coarse for unambiguous classification. Obligate cave-dwellers could be a major component of cave biodiversity in relation to cave entrance elevation, internal air temperature and geographic location (Kováč et al., 2016), and they show strong spatial variation in hypogean habitat in relationships with variation of depth, temperature and availability of organic matter (Rendoš et al., 2016, 2020). Future studies could consider a functional analysis of traits of detected MOTUs, including traits that can undergo troglomorphic changes (such as the modification of the foot complex and elongation of appendages; Christiansen, 1965; Thibaud and Deharveng, 1994; Kováč et al., 2016). More complete reference databases are needed for such a detailed analysis.

5. Conclusion

The pattern of eDNA detection in cave soils was unexpectedly complex, suggesting that cave soil eDNA comprises multiple typologies of eDNA, including local eDNA from soil organisms, probably representing a fine-scale ecological signal, and allochthonous eDNA from the surrounding landscape. So far, the analysis of eDNA detection probability has mostly been used to evaluate species presence and to estimate occupancy probability across surveyed sites. Our study supports the idea that analyzing eDNA detection probability can provide unexploited information on the complex processes underlying eDNA detection (Furlan et al., 2016; Song et al., 2017; Troth et al., 2021). For instance, we suggest that taxa with high detection probability and low occupancy most likely represent local assemblage and could provide direct information on the understudied subterranean cave organisms, while the ones with low detection probability and high occupancy could more often represent allochthonous eDNA. Data on detection probability can also be combined with life history information for a more detailed picture of ongoing processes. Further studies, comparing the outcomes of different approaches, are required to test this hypothesis.

Author contribution

EL collected the samples, AG, AB and GFF performed analyses. GFF drafted the manuscript and prepared tables and figures, all authors contributed in revising the manuscript and accepted its final version.

Data and materials availability

All relevant data have been submitted as supplementary files.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This study was partially funded by the European Research Council under the European Community Horizon 2020 Program, Grant Agreement no. 772284 (IceCommunities). F.C. received a partial funding from the Ministry of University and Research of Italy (MUR), project FOE 2020-Capitale naturale e risorse per il futuro dell'Italia-Task Biodiversità.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.scitotenv.2022.154022>.

References

- Alsos, I.G., Lammers, Y., Yoccoz, N.G., Jørgensen, T., Sjögren, P., Gielly, L., et al., 2018. Plant DNA metabarcoding of lake sediments: how does it represent the contemporary vegetation. *PLoS ONE* 13, e0195403.
- Andersen, K., Bird, K.L., Rasmussen, M., Haile, J., Breuning-Madsen, H., Kjær, K.H., et al., 2012. Meta-barcoding of 'dirty' DNA from soil reflects vertebrate biodiversity. *Mol. Ecol.* 21, 1966–1979.
- Balian, E.V., Lévêque, C., Segers, H., Martens, K., 2007. *Freshwater Animal Diversity Assessment*. Developments in Hydrobiology. Springer, Dordrecht, pp. 329–338.
- Barzaghi, B., Ficetola, G.F., Pennati, R., Manenti, R., 2017. Biphase predators provide biomass subsidies in small freshwater habitats: a case study of spring and cave pools. *Freshw. Biol.* 62, 1637–1644.
- Bates, D., Maechler, M., Bolker, B.M., Walker, S., 2015. Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* 67, 1–48.
- Bauernfeind, E., Soldan, T., 2012. *The Mayflies of Europe (Ephemeroptera)*. Apollo books, Ollerup.
- Bienert, F., De Danieli, S., Miquel, C., Coissac, E., Poillot, C., Brun, J.J., et al., 2012. Tracking earthworm communities from soil DNA. *Mol. Ecol.* 21, 2017–2030.
- Bonin, A., Guerrieri, A., Ficetola, G.F., 2021. Optimal sequence similarity thresholds for clustering of molecular operational taxonomic units in DNA metabarcoding studies. *Authorea* <https://doi.org/10.22541/au.163398436.69595421/v1>.
- Boyd, S.H., Niemiller, K.D.K., Dooley, K.E., Nix, J., Niemiller, M.L., 2020. Using environmental DNA methods to survey for rare groundwater fauna: detection of an endangered endemic cave crayfish in northern Alabama. *PLoS One* 15, e0242741.
- Boyer, F., Mercier, C., Bonin, A., Le Bras, Y., Taberlet, P., Coissac, E., 2016. Obitools: a unix-inspired software package for DNA metabarcoding. *Mol. Ecol. Resour.* 16, 176–182.
- Calderón-Sanou, I., Münkemüller, T., Boyer, F., Zinger, L., Thuiller, W., 2020. From environmental DNA sequences to ecological conclusions: how strong is the influence of methodological choices? *J. Biogeogr.* 47, 193–206.
- Canedoli, C., Ficetola, G.F., Corengia, D., Tognini, P., Ferrario, A., Padoa-Schioppa, E., 2022. Integrating landscape ecology and the assessment of ecosystem services in the study of karst areas. *Landsc. Ecol.* 37, 347–365. <https://doi.org/10.1007/s10980-021-01351-2>.
- Capo, E., Giguet-Covex, C., Rouillard, A., Nota, K., Heintzman, P., Vuillemin, A., et al., 2021. Lake sedimentary DNA research on past terrestrial and aquatic biodiversity: overview and recommendations. *Quaternary* 4, 6. <https://doi.org/10.3390/quat4010006>.
- Chen, W., Ficetola, G.F., 2019. Conditionally autoregressive models improve occupancy analyses of autocorrelated data: an example with environmental DNA. *Mol. Ecol. Resour.* 19, 163–175.
- Chen, W., Ficetola, G.F., 2020. Statistical and numerical methods for sedimentary-ancient-DNA-based study on past biodiversity and ecosystem functioning. *Environ. DNA* 2, 115–129.
- Christiansen, K., 1965. Behavior and form in the evolution of cave collembola. *Evolution* 19, 529–537.
- Clarke, L.J., Soubrier, J., Weyrich, L.S., Cooper, A., 2014. Environmental metabarcodes for insects: in silico PCR reveals potential for taxonomic bias. *Mol. Ecol. Resour.* 14, 1160–1170.
- Culver, D.C., Pipan, T., 2009. *The Biology of Caves and Other Subterranean Habitats*. Oxford University Press, Oxford.
- Culver, D.C., Trontelj, P., Zagmajster, M., Pipan, T., 2012. Paving the way for standardized and comparable subterranean biodiversity studies. *Subterranean Biol.* 10, 43–50.
- Dallai, R., Malatesta, E., 1982. *Collemboli Cavemicoli Italiani (Ricerche sui collemboli XXVI)*. Lavori della Società Italiana di Biogeografia. 7, pp. 173–194.
- Deharveng, L., Bedos, A., 2018. Diversity of terrestrial invertebrates in subterranean habitats. In: Moldovan, O.T., Kováč, L., Halse, S. (Eds.), *Cave Ecology*. Springer, Berlin, pp. 107–172.
- Deiner, K., Fronhofer, E.A., Machler, E., Wasler, J.C., Altermatt, F., 2016. Environmental DNA reveals that rivers are conveyor belts of biodiversity information. *Nat. Commun.* 7, 12544.
- Evans, N.T., Olds, B.P., Renshaw, M.A., Turner, C.R., Li, Y., Jerde, C.L., et al., 2016. Quantification of mesocosm fish and amphibian species diversity via environmental DNA metabarcoding. *Mol. Ecol. Resour.* 16, 29–41.
- Ficetola, G.F., Pansu, J., Bonin, A., Coissac, E., Giguet-Covex, C., De Barba, M., et al., 2015. Replication levels, false presences, and the estimation of presence / absence from eDNA metabarcoding data. *Mol. Ecol. Resour.* 15, 543–556.
- Ficetola, G.F., Taberlet, P., Coissac, E., 2016. How to limit false positives in environmental DNA and metabarcoding? *Mol. Ecol. Resour.* 16, 604–607.
- Ficetola, G.F., Canedoli, C., Stoch, F., 2019a. The racovitzan impediment and the hidden diversity of unexplored environments. *Conserv. Biol.* 33, 214–216.
- Ficetola, G.F., Manenti, R., Taberlet, P., 2019b. Environmental DNA and metabarcoding for the study of amphibians and reptiles: species distribution, the microbiome, and much more. *Amphibia-Reptilia* 40, 129–148.
- Ficetola, G.F., Bonin, A., Boyer, F., Valentini, A., Meyer, A., Dejean, T., et al., 2021. Comparison of markers for the monitoring of freshwater benthic biodiversity through DNA metabarcoding. *Mol. Ecol.* 30, 3189–3202.
- Fierer, N., Strickland, M.S., Liptzin, D., Bradford, M.A., Cleveland, C.C., 2009. Global patterns in belowground communities. *Ecol. Lett.* 12, 1238–1249.
- Fiser, C., Pipan, T., Culver, D.C., 2014. The vertical extent of groundwater metazoans: an ecological and evolutionary perspective. *Bioscience* 64, 971–979.
- Fonseca, V.G., 2018. Pitfalls in relative abundance estimation using eDNA metabarcoding. *Mol. Ecol. Resour.* 18, 923–926.
- Furlan, E.M., Gleeson, D., Hardy, C.M., Duncan, R.P., 2016. A framework for estimating the sensitivity of eDNA surveys. *Mol. Ecol. Resour.* 16, 641–654.
- Gibert, J., Culver, D.C., 2009. Assessing and conserving groundwater biodiversity: an introduction. *Freshw. Biol.* 54, 639–648.
- Gibert, J., Deharveng, L., 2002. Subterranean ecosystems: a truncated functional biodiversity. *Bioscience* 52, 473–481.
- Giguet-Covex, C., Ficetola, G.F., Walsh, K., Poulenard, J., Bajard, M., Fouinat, L., et al., 2019. New insights on lake sediment DNA from the catchment: importance of taphonomic and analytical issues on the record quality. *Sci. Rep.* 9, 14676.
- Goricki, S., Stankovic, D., Snoj, A., Kuntner, M., Jeffery, W.R., Trontelj, P., et al., 2017. Environmental DNA in subterranean biology: range extension and taxonomic implications for proteus. *Sci. Rep.* 7, 45054.
- Guerrieri, A., Bonin, A., Münkemüller, T., Gielly, L., Thuiller, W., Ficetola, G.F., 2021. Effects of soil preservation for biodiversity monitoring using environmental DNA. *Mol. Ecol.* 30, 3313–3325.
- Guillera-Arroita, G., 2017. Modelling of species distributions, range dynamics and communities under imperfect detection: advances, challenges and opportunities. *Ecography* 40, 281–295.
- Guillera-Arroita, G., Lahoz-Monfort, J.J., van Rooyen, A.R., Weeks, A.R., Tingley, R., 2017. Dealing with false-positive and false-negative errors about species occurrence at multiple levels. *Methods Ecol. Evol.* 8, 1081–1091.
- Haile, J., Holdaway, R., Oliver, K., Bunce, M., Gilbert, M.T.P., Nielsen, R., et al., 2007. Ancient DNA chronology within sediment deposits: are paleobiological reconstructions possible and is DNA leaching a factor? *Mol. Biol. Evol.* 24, 982–989.
- Hopkin, S.P., 1997. *Biology of the Springtails (Insecta: Collembola)*. Oxford University Press, Oxford.
- Howarth, F.G., Moldovan, O.T., 2018. The ecological classification of cave animals and their adaptations. In: Moldovan, O.T., Kováč, L., Halse, S. (Eds.), *Cave Ecology*. Springer, Berlin, pp. 41–46.
- Janssen, P., Bec, S., Fuhr, M., Taberlet, P., Brun, J.-J., Bouget, C., 2018. Present conditions may mediate the legacy effect of past land-use changes on species richness and composition of above- and below-ground assemblages. *J. Ecol.* 106, 306–318.
- Kováč, L., Parimuchová, A., Miklisová, D., 2016. Distributional patterns of cave collembola (Hexapoda) in association with habitat conditions, geography and subterranean refugia in the Western Carpathians. *Biol. J. Linn. Soc.* 119, 571–592.
- Kuznetsova, A., Brockhoff, B., Christensen, H.B., 2017. lmerTest package: tests in linear mixed effects models. *J. Stat. Softw.* 82, 1–26.
- Lahoz-Monfort, J.J., Guillera-Arroita, G., Tingley, R., 2016. Statistical approaches to account for false positive errors in environmental DNA samples. *Mol. Ecol. Resour.* 16, 673–685.
- Lunghi, E., Manenti, R., Ficetola, G.F., 2014. Do cave features affect underground habitat exploitation by non-troglobite species? *Acta Oecol.* 55, 29–35.
- Lunghi, E., Manenti, R., Ficetola, G.F., 2017. Cave features, seasonality and subterranean distribution of non-obligate cave dwellers. *PeerJ* 5, e3169. <https://doi.org/10.7717/peerj.3169>.
- Lunghi, E., Manenti, R., Mulargia, M., Veith, M., Corti, C., Ficetola, G.F., 2018. Environmental suitability models predict population density, performance and body condition for microendemic salamanders. *Sci. Rep.* 8, 7527.
- Lunghi, E., Ficetola, G.F., Zhao, Y., Manenti, R., 2020. Are the neglected tipuloidea crane flies (Diptera) an important component for subterranean environments? *Diversity* 12, 333.
- Maaß, S., Caruso, T., Rillig, M.C., 2015. Functional role of microarthropods in soil aggregation. *Pedobiologia* 58, 59–63.
- Mächler, E., Walser, J.-C., Altermatt, F., 2021. Decision-making and best practices for taxonomy-free environmental DNA metabarcoding in biomonitoring using hill numbers. *Mol. Ecol.* 30, 3326–3339.
- MacNeil, R.R., Brcic, J., 2017. Coping with the subterranean environment: a thematic content analysis of the narratives of cave explorers. *J. Hum. Perform. Extreme Environ.* 13, 6.
- Mammola, S., 2019. Finding answers in the dark: caves as models in ecology fifty years after poulsen and white. *Ecography* 42, 1331–1351.
- Mammola, S., Amorim, I.R., Bichuette, M.E., Borges, P.A.V., Cheeptham, N., Cooper, S.J.B., et al., 2020. Fundamental research questions in subterranean biology. *Biol. Rev.* 95, 1855–1872.
- Mammola, S., Lunghi, E., Bilandžija, H., Cardoso, P., Grimm, V., Schmidt, S.I., et al., 2021. Collecting eco-evolutionary data in the dark: impediments to subterranean research and how to overcome them. *Ecol. Evol.* 11, 5911–5926.

- Martins, F.M.S., Porto, M., Feio, M.J., Egeter, B., Bonin, A., Serra, S.R.Q., et al., 2021. Modelling technical and biological biases in macroinvertebrate community assessment from bulk preservative using multiple metabarcoding markers. *Mol. Ecol.* 30, 3221–3238.
- Mercier, C., Boyer, F., Bonin, A., Coissac, E., 2013. SUMATRA and SUMACLUSt: fast and exact comparison and clustering of sequences. Programs and Abstracts of the SeqBio 2013 Workshop (Abstract). GdRBIM and gdrIM, Montpellier, France, pp. 27–29.
- Minelli, A., Stoch, F., 2006. The checklist of the Italian fauna. *Mem. Museo Civico Storia Nat. Verona* 2007 (17), 21–23.
- Nichols, R.V., Vollmers, C., Newsom, L.A., Wang, Y., Heintzman, P.D., Leighton, M., et al., 2018. Minimizing polymerase biases in metabarcoding. *Mol. Ecol. Resour.* 18, 927–939.
- Pansu, J., Winkworth, R.C., Hennion, F., Gielly, L., Taberlet, P., Choler, P., 2015. Long-lasting modification of soil fungal diversity associated with the introduction of rabbits to a remote sub-Antarctic archipelago. *Biol. Lett.* 11, 20150408.
- Parducci, L., Bennet, K.D., Ficetola, G.F., Alsos, I.G., Suyama, Y., Wood, J.R., et al., 2017. Ancient plant DNA in lake sediments. *New Phytol.* 214, 924–942.
- Popovic, S., Krizmanic, J., Vidakovic, D., Karadzic, V., Milovanovic, Z., Pecic, M., et al., 2020. Biofilms in caves: easy method for the assessment of dominant phototrophic groups/taxa in situ. *Environ. Monit. Assess.* 192.
- Potapov, A.A., Semenina, E.E., Korotkevich, A.Y., Kuznetsova, N.A., Tiunov, A.V., 2016. Connecting taxonomy and ecology: trophic niches of collembolans as related to taxonomic identity and life forms. *Soil Biol. Biochem.* 101, 20–31.
- Rendoš, M., Miklisová, D., Kováč, L., Mock, A., 2020. Dynamics of Collembola (Hexapoda) in a forested limestone scree slope, Western Carpathians, Slovakia. *J. Cave Karst Studies* 82 (1), 18–29. <https://doi.org/10.4311/2018LSC0140>.
- Rendoš, M., Raschmanova, N., Kovac, L., Miklisova, D., Mock, A., Luptacik, P., 2016. Organic carbon content and temperature as substantial factors affecting diversity and vertical distribution of collembola on forested scree slopes. *Eur. J. Soil Biol.* 75.
- Romero, A., 2009. *Cave Biology: Life in Darkness*. Cambridge University Press, Cambridge.
- Rosero, P., Crespo-Pérez, V., Espinosa, R., Andino, P., Barragán, A., Moret, P., et al., 2021. Multi-taxa colonisation along the foreland of a vanishing equatorial glacier. *Ecography* 44, 1010–1021.
- Rota, N., Canedoli, C., Ferrè, C., Ficetola, G.F., Guerrieri, A., Padoa-Schioppa, E., 2020. Evaluation of Soil Biodiversity in Alpine Habitats Through eDNA Metabarcoding and Relationships with Environmental Features Forests. 11, p. 738.
- Saccò, M., Blyth, A., Bateman, P.W., Hua, Q., Mazumder, D., White, N., et al., 2019. New light in the dark - a proposed multidisciplinary framework for studying functional ecology of groundwater fauna. *Sci. Total Environ.* 662, 963–977.
- Salgado, S.S., Motta, P.C., Aguiar, L.M.D., Nardoto, G.B., 2014. Tracking dietary habits of cave arthropods associated with deposits of hematophagous bat guano: a study from a neotropical savanna. *Austral Ecol.* 39, 560–566.
- Schneider, K., Christman, M.C., Fagan, W.F., 2011. The influence of resource subsidies on cave invertebrates: results from an ecosystem-level manipulation experiment. *Ecology* 92, 765–776.
- Slon, V., Hopfe, C., Weiß, C.L., Mafessoni, F., de la Rasilla, M., Lalueza-Fox, C., et al., 2017. Neandertal and denisovan DNA from pleistocene sediments. *Science* 356, 605.
- Smithson, M., Verkuilen, J., 2006. A better lemon squeezer? Maximum-likelihood regression with beta-distributed dependent variables. *Psychol. Methods* 11, 54–71.
- Song, J.W., Small, M.J., Casman, E.A., 2017. Making sense of the noise: the effect of hydrology on silver carp eDNA detection in the Chicago area waterway system. *Sci. Total Environ.* 605, 713–720.
- Su, Y.S., Yajima, M., 2015. R2jags: Using R to Run 'JAGS'. R Package Version 0.5-7: www.r-project.org.
- Taberlet, P., Prud'homme, S.M., Campione, E., Roy, J., Miquel, C., Shehzad, W., et al., 2012. Soil sampling and isolation of extracellular DNA from large amount of starting material suitable for metabarcoding studies. *Mol. Ecol.* 21, 1816–1820.
- Taberlet, P., Bonin, A., Zinger, L., Coissac, E., 2018. Environmental DNA for Biodiversity Research and Monitoring. Oxford University Press, Oxford.
- Thibaud, J.M., Deharveng, L., 1994. Collembola. In: Juberthie, C., Decu, V. (Eds.), *Encyclopaedia Biospeologica. 1. Société de Biospéologie, Moulis*, pp. 267–276.
- Troth, C.R., Sweet, M.J., Nightingale, J., Burian, A., 2021. Seasonality, DNA degradation and spatial heterogeneity as drivers of eDNA detection dynamics. *Sci. Total Environ.* 768.
- Venarsky, M.P., Huntsman, B.M., Huryn, A.D., Benstead, J.P., Kuhajda, B.R., 2014. Quantitative food web analysis supports the energy-limitation hypothesis in cave stream ecosystems. *Oecologia* 176, 859–869.
- Vörös, J., Márton, O., Schmidt, B.R., Gál, J.T., Jelić, D., 2017. Surveying Europe's only cave-dwelling chordate species (Proteus anguinus) using environmental DNA. *PLoS ONE* 12, e0170945.
- West, K.M., Richards, Z.T., Harvey, E.S., Susac, R., Grealy, A., Bunce, M., 2020. Under the karst: detecting hidden subterranean assemblages using eDNA metabarcoding in the caves of Christmas Island, Australia. *Sci. Rep.* 10, 21479.
- White, W.B., 2007. Cave sediments and paleoclimate. *J. Cave Karst Stud.* 69, 76–93.
- White, N.E., Guzik, M.T., Austin, A.D., Moore, G.I., Humphreys, W.F., Alexander, J., et al., 2020. Detection of the rare Australian endemic blind cave eel (*Ophisternon candidum*) with environmental DNA: implications for threatened species management in subterranean environments. *Hydrobiologia* 847, 3201–3211.
- Yoccoz, N.G., Brathen, K.A., Gielly, L., Haile, J., Edwards, M.E., Goslar, T., et al., 2012. DNA from soil mirrors plant taxonomic and growth form diversity. *Mol. Ecol.* 21, 3647–3655.
- Zhang, D., Xia, H., Chen, F., Li, B., Slon, V., Cheng, T., et al., 2020. Denisovan DNA in late pleistocene sediments from baishiya karst cave on the tibetan plateau. *Science* 370, 584.
- Zinger, L., Bonin, A., Alsos, I., Bálint, M., Bik, H., Boyer, F., et al., 2019a. DNA metabarcoding - need for robust experimental designs to draw sound ecological conclusions. *Mol. Ecol.* 28, 1857–1862.
- Zinger, L., Taberlet, P., Schimann, H., Bonin, A., Boyer, F., De Barba, M., et al., 2019b. Body size determines soil community assembly in a tropical forest. *Mol. Ecol.* 28, 528–543.