Received: 28 November 2022 | Revised: 31 March 2023 | Accepted: 5 April 2023

DOI: 10.1111/mec.16955

ORIGINAL ARTICLE

MOLECULAR ECOLOGY WILEY

Groundwater environmental DNA metabarcoding reveals hidden diversity and reflects land-use and geology

Marjorie Couton[1] | Samuel Hürlemann[1] | Angela Studer[1] | Roman Alther[1,2] | Florian Altermatt[1,2]

[1]Department of Aquatic Ecology, Eawag: Swiss Federal Institute of Aquatic Science and Technology, Dübendorf, Switzerland

[2]Department of Evolutionary Biology and Environmental Studies, University of Zurich, Zürich, Switzerland

Correspondence
Marjorie Couton and Florian Altermatt, Eawag: Swiss Federal Institute of Aquatic Science and Technology, Department of Aquatic Ecology, Überlandstrasse 133, CH-8600 Dübendorf, Switzerland.
Email: marjorie.couton@eawag.ch and florian.altermatt@eawag.ch

Funding information
Bundesamt für Umwelt; Fachverband für Gas, Wärme und Wasser (SVGW/SSIGE); Schweizerischer Nationalfonds zur Förderung der Wissenschaftlichen Forschung, Grant/Award Number: 31003A_173074; Universität Zürich

Handling Editor: Jacob A Russell

Abstract

Despite being the most important source of liquid freshwater on the planet, groundwater is severely threatened by climate change, agriculture, or industrial mining. It is thus extensively monitored for pollutants and declines in quantity. The organisms living in groundwater, however, are rarely the target of surveillance programmes and little is known about the fauna inhabiting underground habitats. The difficulties accessing groundwater, the lack of expertise, and the apparent scarcity of these organisms challenge sampling and prohibit adequate knowledge on groundwater fauna. Environmental DNA (eDNA) metabarcoding provides an approach to overcome these limitations but is largely unexplored. Here, we sampled water in 20 communal spring catchment boxes used for drinking water provisioning in Switzerland, with a high level of replication at both filtration and amplification steps. We sequenced a portion of the COI mitochondrial gene, which resulted in 4917 ASVs, yet only 3% of the reads could be assigned to a species, genus, or family with more than 90% identity. Careful evaluation of the unassigned reads corroborated that these sequences were true COI sequences belonging mostly to diverse eukaryotic groups, not present in the reference databases. Principal component analyses showed a strong correlation of the community composition with the surface land-use (agriculture vs. forest) and geology (fissured rock vs. unconsolidated sediment). While incomplete reference databases limit the assignment of taxa in groundwater eDNA metabarcoding, we showed that taxonomy-free approaches can reveal large hidden diversity and couple it with major land-use drivers, revealing their imprint on chemical and biological properties of groundwater.

KEYWORDS
biodiversity, eDNA, metazoans, reference databases, stygofauna, taxonomic assignment

1 | INTRODUCTION

Groundwater is the largest freshwater reservoir on earth (96%, excluding glaciers and ice caps), providing drinking water for half of the world's population (Smith et al., 2016). It also plays a major role in maintaining riparian and wetland environments during dry periods or in the context of climate change (Somers et al., 2019). Groundwater provides invaluable services but is also facing many threats, including depletion from overuse and pollution from industry and agriculture (Burri et al., 2019). Consequently, groundwater quantity and quality are extensively monitored. Many national and international management plans, and their respective monitoring, aim to control its consumption for a sustainable use, and to limit pollutants reaching the aquifers (e.g. the Groundwater Directive (2006/118/EC) from EU's Water Framework Directive, or the Sustainable Groundwater Management Act in California). Among all these regulations, however, none is considering groundwater as a habitat for subterranean organisms. Even if the general recommendations for biodiversity protection can be applied to any type of environment (Sutherland et al., 2018), no specific policies have been directed towards the study and preservation of underground organisms.

Groundwater organisms provide useful ecosystem services for the quality of ground and surface waters (Mammola et al., 2019). For example, micro-organisms can transform and degrade contaminants, and invertebrates can maintain hydraulic conductivity through their feeding on biofilms and bioturbation (Boulton et al., 2008; Griebler & Avramov, 2015). The latter (and particularly amphipods) are commonly used as bioindicators in monitoring surface or groundwater ecosystems (e.g. Koch et al., 2021). Studying these organisms, however, is a major challenge, because most subterranean voids are not or very poorly accessible to humans, and much of the groundwater fauna is insufficiently known from a taxonomic perspective. Even though the number of underground species described is ever increasing since the last century (see example of karst cave species records in Ficetola et al., 2019; Zagmajster et al., 2018), species richness and community composition are substantially underestimated. Moreover, our knowledge of described species' distribution is very partial (Deharveng et al., 2009). There is thus a dire need for a more extensive exploration of groundwater habitats.

Environmental DNA (eDNA) is being more and more frequently used for the study and monitoring of surface aquatic environments (Deiner et al., 2017; Keck et al., 2022; Taberlet et al., 2012). Sampling eDNA is minimally destructive and offers the possibility to study rare and/or elusive species (e.g. Mächler et al., 2014; Pfleger et al., 2016; Tang et al., 2019). Applying a metabarcoding approach on eDNA samples allows the identification of a high number of taxa without the need for taxonomical expertise, a skill that can be particularly scarce for underground organisms. Moreover, this technique can target a broad set of taxa with only one sample. For all these reasons, eDNA can be a valuable tool in the monitoring of groundwater species.

A handful of studies have started to collect water and soil samples in subterranean environments for the use of eDNA-based approaches (Saccò et al., 2022 and references therein). However, the focus has been largely on the detection of particular individual taxa using species-specific techniques. For example, Niemiller et al. (2018) detected species from the amphipod genus *Stygobromus* in cave lakes, using qPCR. Further studies explored the use of eDNA metabarcoding to evaluate groundwater biodiversity in aquifers, either targeting solely prokaryotes (e.g. Korbel et al., 2022; Morse et al., 2021; Voisin et al., 2020), or also including micro-eukaryotic communities (e.g. Herrmann et al., 2020). Only a few studies, however, focused on macro-eukaryotes (Korbel et al., 2017; Oberprieler et al., 2021; West et al., 2020), with contrasting results. West et al. (2020) were able to identify 60 different orders of metazoans with eDNA from Australian caves, and Oberprieler et al. (2021) detected signals of underground taxa in bore holes for which no actual organisms were observed. Yet, some key organisms, such as amphipods or syncarids that were observed with net samples, were not retrieved with eDNA in Korbel et al. (2017). Moreover, all three of these studies pinpoint the low taxonomic resolution of metabarcoding assignments due to a lack of references in public databases.

Overall, there is a need for a thorough evaluation of the possibilities and limitations of eDNA metabarcoding methods in the context of studying underground organisms. Since biodiversity assessments are generally not available for groundwater habitats, a direct comparison between eDNA metabarcoding and traditional sampling – the standard calibration for eDNA data in other environments (Keck et al., 2022) – is not possible. Furthermore, groundwater biodiversity being generally understudied, large gaps in the reference databases may hinder the understanding of such samples.

Here, we developed an extensive protocol based on eDNA metabarcoding to study the metazoan diversity living in aquifers. Groundwater samples were collected in 20 spring catchment boxes (SCBs; i.e. passive water collectors from a shallow aquifer using underground drainage pipes; see Figure 1b,c), with the participation of drinking water providers in North-Eastern Switzerland. We used a high number of replicates and large volume of water filtered (compared to aboveground freshwater standards) to account for the potentially low density of metazoan organisms in groundwater. Following DNA extraction, we amplified a portion of the COI mitochondrial gene and evaluated our ability to retrieve diverse types of metazoans as well as the suitability of our sampling design. Since taxonomic assignments are difficult for underground taxa due to the lack of references, we carefully checked the unassigned portion of our data set to make sure that it could be used for diversity assessments.

2 | MATERIALS AND METHODS

2.1 | Sampling

We collected the samples in the catchment basin of the river Töss in North-Eastern Switzerland (Figure 1a). The 20 sampled sites are located in the Swiss Molasse Basin, containing a mix of

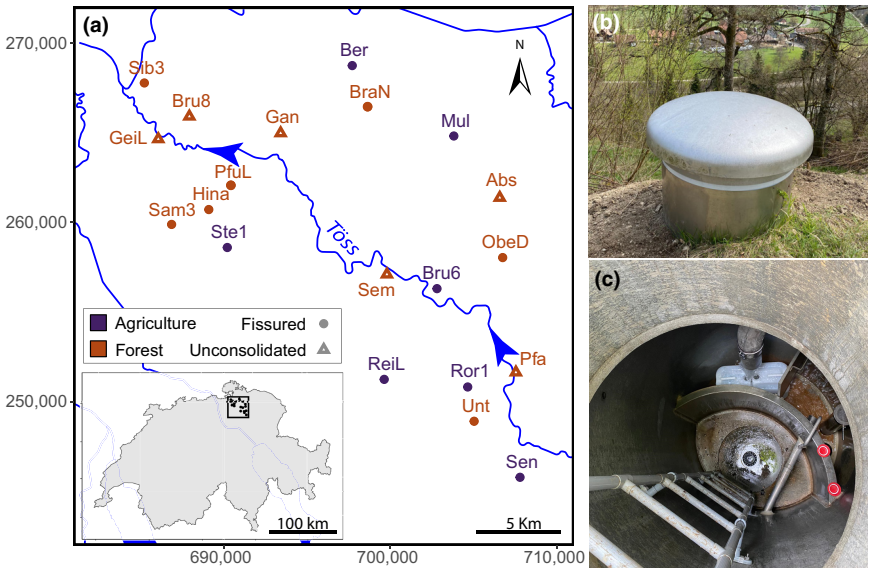

FIGURE 1 (a) Map of the Töss catchment (NE Switzerland), displaying the 20 sampled spring catchment boxes (SCB). The colours indicate the land-use surrounding the sampling site: orange for forest and purple for agriculture. Filled circles indicate that the water sampled comes from a fissured aquifer whereas open triangles indicate that it comes from an unconsolidated aquifer. The coordinates used are from the Swiss LV03 system. (b) and (c) are illustrating a spring catchment box from the outside and inside, respectively.

unconsolidated and fissured aquifers. Groundwater is the primary source of drinking water in this area, and we gathered the water samples in SCBs (Figure 1b,c) in collaboration with water providers. The sampled aquifers are rather shallow, and the collection pipes are emerging between 1 and 5 m below ground. All water samples were passively collected (i.e. from the natural flow of water coming out of the aquifer) and no pumping stations were included. Each location was categorized depending on the type of aquifer from which the water was collected (derived from swisstopo's groundwater bodies geodata, https://map.geo.admin.ch) and with respect to the land-use surrounding the SCB (Table S1). Each site was classified as either forested area or agricultural area (incl. cultivated cropping fields, pastures, and a golf course) based on a visual inspection at the time of sampling. All sites located at a forest edge were put in the forest category.

We conducted the sampling between April and June 2021. At each site, we sampled 40 L of water in canisters that were brought back to the laboratory in cooling boxes within 2 h for subsequent filtration. The water was collected directly from the outflow pipe of the SCB to make sure that the samples were representative of the aquifer itself. We used enclosed Sterivex™ filter units (Merck Millipore) to filter 4 × 10 L per site. The filter units are made of polyethersulfone, with a pore size of 0.22 μm and a filtration area of 10 cm^2. We used a peristaltic pump (Alexis®; Proactive Environmental Products) with a flow rate of approximately 6.7 mL/s. All filter units were then stored at −20°C until DNA extraction. We wore gloves at all sampling steps, including material preparation, to reduce contamination. We cleaned and reused the canisters and tubing for the pump between sites by soaking them into a bleach solution (2 L of commercial <2% hypochlorite solution diluted with 3 L of molecular grade water) for at least 30 min and then rinsed three times with Nanopure™ water. We performed three filter controls by collecting Nanopure™ water in the canisters following the decontamination protocol detailed above. Ten litres of this water was filtered for each control following the same protocol as the other samples.

2.2 | DNA extraction and library construction

We performed DNA extractions and the first PCR step in a clean lab with constant air overpressure and no presence of PCR products to reduce the risk of DNA contamination. Cleanliness protocols and safety procedures follow previously established uses (Brantschen et al., 2021; Deiner et al., 2015; Pawlowski et al., 2020). We extracted DNA using the DNeasy® PowerWater® Sterivex™ Kit (Qiagen®) following the manufacturer's protocol, except for the bead-beating steps (steps 12 and 13) that were omitted. We preheated the elution buffer at 70°C, and passed it through the column twice in order to increase DNA yield. We produced three extraction controls by adding 0.9 mL ST1 buffer to clean Sterivex™ filter units, and by processing them following the same protocol as the other samples. All samples were quantified using a Nanodrop™ spectrophotometer (Thermo Fisher Scientific Inc.), yet the amount of DNA was below the threshold of detection of the device for many of them (16 samples). We thus diluted all samples at 1/10th of their original concentration (ranging from <2 to 30 ng/μL, with a mean concentration of 7.8 ng/μL), considering that they were all in the same order of magnitude.

Based on previous knowledge (Studer, 2022), amphipod species from the genus *Niphargus* occur at a high abundance and diversity in groundwater at our sampling sites. However, DNA fragments from these species are generally not well-amplified by primers traditionally used in metabarcoding (e.g. Brantschen et al., 2022; Elbrecht & Leese, 2017; Leray et al., 2013). Thus, we modified the primers developed by Vamos et al. (2017), amplifying a short fragment of the COI gene (205 bp), as follows: fwhF2_Niph 5′-GGRTGAACAGT WTAYCCTCC-3′ and fwhR2n_Niph 5′-GTRATWGCTCCWGCTAR MACTGG-3′. The range of amplification of these primers was tested by an in silico PCR on GenBank's nt database, using the ecoPCR command from the OBITOOLS-v1.2.11 package (Boyer et al., 2016; Figure S1). We also evaluated their ability to amplify *Niphargus* sp. DNA on individual DNA samples from 48 different *Niphargus* species.

The library preparation was conducted in a 2-step PCR approach. DNA from each filter was amplified in five separate PCR replicates, where each fragment was given a unique 8-bp tag combination at both ends, in order to distinguish between them (see Table S2 for details). Each tagged replicate was composed of a pool of three individual PCRs performed in a total volume of 10 μL. We started with a PCR mix of a volume of 30 μL, consisting of 15 μL of KAPA® HiFi HotStart ReadyMix, 2 μL of each forward and reverse primer (10 μM original concentration), 3 μL of extracted DNA, and 8 μL of molecular grade water. We split the total volume of 30 μL reaction mix into 3 individual plates (10 μL each) and distributed them in 3 independent Thermocyclers (Biometra T1Thermocycler, AnalytikJena GmbH). PCR cycling started by a denaturation step at 95°C for 5 min, followed by 35 cycles of 98°C for 30 s, 48°C for 30 s and 72°C for 1 min. The final extension step lasted 10 min at 72°C. We then pooled the three 10 μL PCR products back together. This step included two PCR controls. We checked the amplification success with a QiAxcel® Screening Cartridge (Qiagen®) and then pooled all five PCR replicates from a sample together according to the intensity of the QiAxcel® band. To reduce the amount of unused primers in the solution, we performed a bead cleaning step using KAPA® pure beads (KAPA Biosystems) with a 0.8 ratio and following the manufacturer's protocol.

A second PCR step was performed to bind the Nextera® index adapters (Set A) to the fragments. Each 30 μL indexing reaction consisted of 15 μL of KAPA® HiFi HotStart ReadyMix, 3 μL of each forward and reverse Nextera® primers, 6 μL of molecular grade water, and 3 μL of cleaned up PCR product from the first PCR step. Analogous to the first step, we split the reaction mix into three independent plates and pooled them back thereafter. PCR cycling started by a denaturation step at 95°C for 5 min, followed by 12 cycles of 98°C for 30 s, 55°C for 30 s and 72°C for 1 min. The final extension step lasted 10 min at 72°C. We then cleaned the indexed amplicons using KAPA® pure beads with a 0.8 ratio and following the manufacturer's protocol.

All samples were pooled, including all controls, at equimolar concentration based on DNA quantification by a selective fluorescence dye Qubit™ BR DNA Assay Kit (Invitrogen™, Thermo Fisher scientific), read on a Spark® Multimode Microplate Reader (Tecan). We cleaned the library pool a last time by using Agencourt® AMPure® XP beads (Beckman Coulter™) with a 0.8 ratio, following the manufacturer's protocol.

The quality of the library was assessed on an Agilent 4200 TapeStation System (Agilent Technologies, Inc.), and quantified its concentration using the Qubit™ (1.0) fluorometer following the manufacturer's protocol for the dsDNA HS Assay. Finally, we performed a paired-end sequencing on an Illumina MiSeq platform (Illumina, Inc.) at the Genomic Diversity Center, ETH Zurich, Switzerland, using a Reagent kit v2 (250 cycles) with 10% PhiX control.

2.3 | Sequencing reads processing

The sequencing device performed the first step, aiming at demultiplexing the samples based on index combination. Then, we assigned reads to each of the five PCR replicates according to their tag combination, and removed primers and tags using CUTADAPT v-2.8 (Martin, 2011). We produced a set of amplicon sequence variants (ASVs) using DADA2 v-1.13.1 (Callahan et al., 2016). To assess 'index-jump' (Taberlet et al., 2018), we added eight unused index combinations to the MiSeq sequencing sample sheet in order to get the corresponding fastq files. For each ASV, we divided the number of reads assigned to one of these internal control index combinations by the total number of reads in the complete data set for this same ASV. We then recorded the maximal proportion occurring in an index control and discarded from a sample any ASV that did not account for more than this maximal proportion in this sample. Furthermore, we retained only ASVs found in at least two out of the 20 replicates per site (four filter replicates × five PCR replicates). All reads present in the negative controls (sampling, extraction and PCR) were processed following the same protocol and none of them contained any reads after correction. All parameters used for each tool and the detailed pipeline for reads processing, taxonomic assignment, and downstream analyses can be found online (https://github.com/joarwrie/NiphToess).

We evaluated the efficiency of our sampling strategy by producing ASV accumulation curves using the *specaccum* function from VEGAN-2.5.7 R package (Oksanen et al., 2018).

We also checked the adequacy of our sequencing strategy through the creation of rarefaction curves for each site using the *rarecurve* function from VEGAN-2.5.7.

2.4 | Taxonomic assignment

We performed a first taxonomic assignment by aligning all ASVs against GenBank nt database (Benson et al., 2013) using the BLAST® command line tool (Altschul et al., 1990) with the default value of 500 hits per query. We considered only alignments with at least 99% of query cover to make sure that even references missing one or two bases at the extremities would be considered. Only alignments displaying a minimum identity of 80% were considered as this value is commonly used as a loose threshold in metabarcoding studies targeting metazoans (e.g. Macher et al., 2018; Mugnai et al., 2023). For each ASV, we identified the alignment with the maximum identity percentage. Then, we assigned this ASV to the lowest common ancestor of all matching references with an identity greater than the maximum identity percentage minus one. For example, if the maximum identity between an ASV and a reference is 99%, all references with more than 98% identity with this ASV will be used for the assignment. We chose to include this 1% identity span to reduce the number of false assignments due to the presence of erroneous sequences in the database. Since it is very unlikely that two haplotypes of a same species diverge more than 10%, we changed all assignments to the species level but with an identity lower than 90% for an assignment to the genus level. Although this threshold is relatively relaxed for species assignment, we believe it is the most appropriate in this particular situation because groundwater organisms are known to harbour many cryptic

species (Lefébure et al., 2006), and because we expect a lack of references from local populations. If an ASV was assigned to references belonging to more than one kingdom, we classified this ASV as 'unassigned'. We checked manually all ASVs assigned to a higher taxonomic level than the species despite having a match in the reference database at a 100% identity for synonymy issues or potential erroneous sequences in the database, and the assignments were modified if necessary. All aquatic metazoan species were then classified according to their affinity to groundwater following the nomenclature of Gibert et al. (1994), using three categories: stygobites, stygophiles, and stygoxenes.

For all ASVs classified as 'unassigned' or assigned to a higher taxonomic level than the family, we performed a second taxonomic assignment using the Dark mAtteR iNvestigator (DARN) tool (Zafeiropoulos et al., 2021). This pipeline is based on a phylogenetic placement algorithm and returns different placement possibilities for each ASV with their likelihood. Our objective was only to identify what type of organisms could be associated with our unknown sequences so we considered only high taxonomic level assignments (kingdom and phylum). We associated each ASV with the taxon with the maximum likelihood weight ratio (LWR), and we only kept this assignment if the LWR was >0.5. If no placement had a LWR >0.5 for a target ASV, we classified it as 'unassigned'.

2.5 | Diversity analyses

We performed all subsequent analyses on the ASV data set, regardless of their assignment status. First, we transformed the data set including all 400 replicates (20 sites×4 filters×5 PCRs) via the 'Hellinger' method (square root of the relative number of reads for each ASV per replicate), as suggested by Legendre and Gallagher (2001), and using the *decostand* function in VEGAN-2.5.7. This transformation accounts for the differences in sequencing cover between the replicates and decreases the weight given to highly abundant ASVs. We visualized a heatmap using GGPLOT2-3.3.6 (Wickham, 2016) on Euclidean distance between pairs of replicates to evaluate the replicability of our approach within sites.

We investigated β-diversity between sites on the ASV data set after pooling all replicates from each site. We again transformed data using the 'Hellinger' method. Then, we visualized the sites through a principal component analysis (PCA) using the *rda* function in VEGAN-2.5.7, as suggested by Borcard et al. (2018). We performed a PERMANOVA (using Euclidean distance) to evaluate the effect of the type of aquifer and land-use on the diversity at each site using the *adonis2* function in VEGAN-2.5.7. As our design is unbalanced (no combination of unconsolidated aquifer and agriculture sites), we used the option to calculate sums of squares sequentially, and we tested alternatively both factors as second term. Because of the 3 months elapsed between the beginning and end of our sampling, we included the month of sampling as a first variable to account for its possible effect on the results.

3 | RESULTS

3.1 | Sequencing results

The MiSeq run yielded 13,584,846 reads with a mean of 139,678 (SD = 104,236) reads per indexed replicate (i.e. at the level of filter unit). The high variation between replicates is congruent with the quality of the PCR amplification, with the one exhibiting the lowest number of reads being the one for which the PCR amplification was poor (no or very faded band on the QIAxcel® electrophoresis). DADA2 processing resulted in 10,799 ASVs, with 4917 (5,657,618 reads) remaining after correction for index-jump and removing sequences present in only one replicate per site. After this step, samples (sites) exhibited a mean number of 282,881 reads (SD = 165,051), and 574 ASVs (SD = 469; Table S3). No reads remained in any of the PCR replicates from the filter number two of the site BraN after above correction steps. All rarefaction curves reached a plateau, indicating that the number of reads for each site were sufficient to retrieve all ASVs present in the samples, even for the samples with the lowest number of reads (Figure 2a).

We evaluated the adequacy of our protocol to represent the genetic diversity at each site by producing ASV accumulation curves (Figure 2b). Accumulation curves at all sites reached a plateau after 15 replicates, whatever the number of ASVs they contained. When looking separately at the effect of filter replicates (all PCR replicates pooled; Figure 2c) and of PCR replicates for each filter (Figure 2d), our results show that four filters were enough to reach a plateau at most sites, except the ones with the highest diversity. For the PCRs, however, the curves are mostly still increasing after five replicates, indicating that our protocol was not completely able to capture the whole diversity present in each filter.

3.2 | Taxonomic assignment

When comparing the sequences against GenBank's nt database, only a small proportion of ASVs (752; 15%), and reads (1,110,324; 20%) were assigned to any taxonomic level with more than 80% identity (Figure 3). When being more stringent and only considering taxonomic assignments with more than 90% identity, an even smaller proportion of ASVs (190; 4%) and reads (178,390; 3%) were identified. Half of the assignments were related to a Protozoan taxon (50%, 465,029 reads, 157 ASVs). The rest was either assigned to Chromista (37%; 343,139 reads, 192 ASVs), Metazoa (7%, 69,517 reads, 142 ASVs), or Fungi (5%, 49,198 reads, 108 ASVs; Figure 4, Table S4). Only 13 ASVs (5116 reads) were assigned to a plant taxon and one ASV (32 reads) to a bacterium taxon. Across these six kingdoms, 62 species were identified (89 ASVs, 21,384 reads), 32 of them being metazoans (Table 1). Despite the loose threshold that we chose for species assignments (90%), only five out of the 32 metazoan species exhibit an identity lower than 97% with their assigned reference. For these ASVs, we need to be

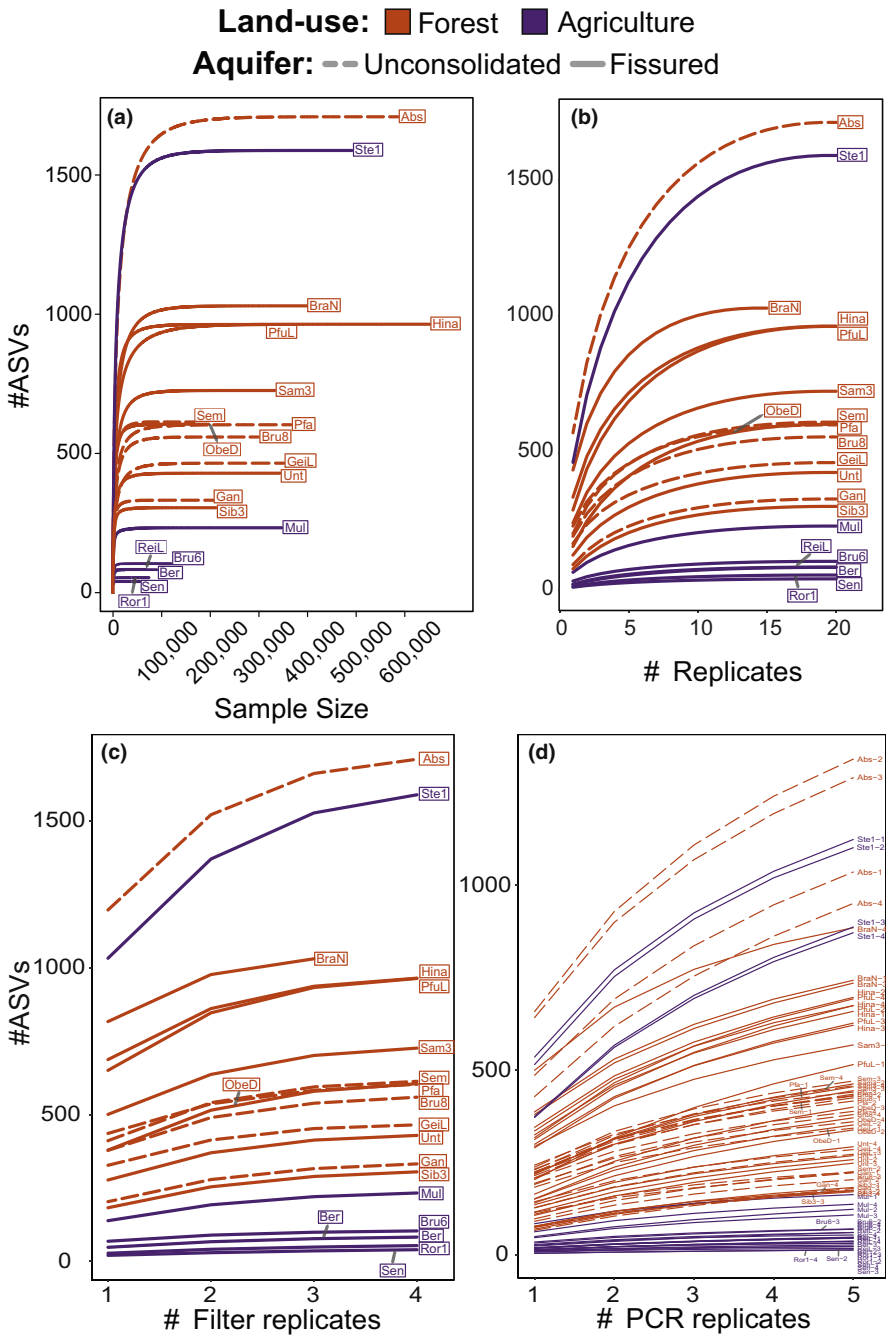

FIGURE 2 (a) Rarefaction curves showing the number of ASVs for an accumulating number of reads randomly picked from the pool available for each site. (b, c) ASV accumulation curves for each site of the study where the number of ASVs is shown for an accumulating number of replicates, includes both filtration and PCR replicates in (b) or including only filter replicates (i.e. all five PCR replicates per filter pooled) in (c). (d) ASV accumulation curves for each filter where the number of ASVs is shown for an accumulating number of PCR replicates. The colours indicate the type of land-use surrounding the sampling site: orange for forest and purple for agriculture. Plain and dashed lines indicate that the water comes from fissured or unconsolidated aquifers, respectively. See Figure 1 for site codes.

careful with our interpretation and keep in mind that they may be erroneous assignments. Out of the 32 metazoan species potentially identified, most of them (19) are arthropods (Table 1). In total, six different phyla are represented, namely, Annelida, Arthropoda, Chordata, Mollusca, Nematoda, and Rotifera. Most of the species are stygoxenes or non-aquatic organisms that have no affinity to the groundwater. Only three are stygophiles (species that can live in the groundwater) and five are stygobites (obligatory groundwater organisms).

All reads unassigned with the BLAST approach, or assigned to a high taxonomic level (order and higher), were processed using a phylogenetic placement method. Most of them could be assigned to a eukaryotic kingdom (75%), and only a small proportion of the unassigned reads were associated with Bacteria (3%) or Archaea (0.01%). The remaining 22% still could not be assigned to any of these groups (Figure 4). Within eukaryotes, 27% of the reads corresponded to Chromista whereas 9% could be attributed to Metazoa and 6% to Protozoa (Figure 4).

3.3 | Diversity analyses

The Euclidean distance between Hellinger-transformed ASV (EHt-ASVs) distributions for each PCR replicate indicates that replicates from the same site are generally more similar to each other than to replicates of different sites (Figure 5). The only exception are the

three sites for which the PCR amplification did not work satisfactorily, that is, Sen, Ror1 and ReiL.

A principal component analysis based on the EHt-ASVs between sites (all replicates pooled) revealed that sites belonging to the same land-use and geology tend to be more similar (Figure 6), except for Ste1, the only site located close to a golf course. The difference between forest and agriculture seems to be the main factor explaining the observed pattern (separation on the first principal component) followed by the aquifer type (separation on the second principal component). The PERMANOVA confirmed the significant impact of land-use with still almost 9% of variance explained after considering the proportion potentially attributed to the aquifer type and the month of sampling (Table 2). The impact of the aquifer type, however, was only significant when included second in the model. Yet, given the clear pattern on the PCA, we can still infer that the aquifer type has an effect on the biodiversity patterns observed. The lack of significance when this term is included last could be explained by a lack of power in our analysis or by a masking effect of land-use. Despite the 3 months span between the first and last sampling date, time of sampling did not have a significant effect. After considering all factors, the variance remaining unexplained is still high (73%).

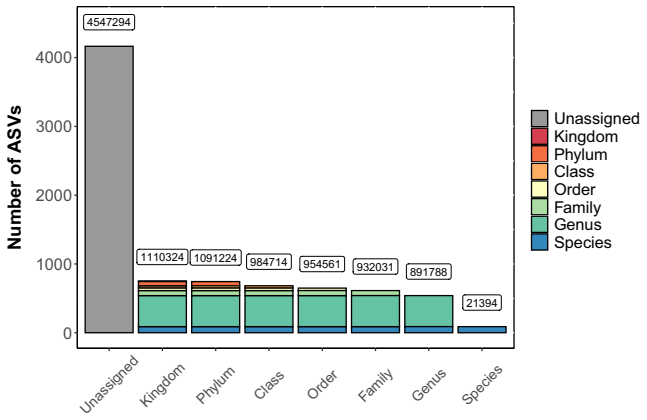

FIGURE 3 Number of ASVs (*y*-axis) and total number of reads (labels on top of each bar) assigned to the different taxonomic levels. The number of ASVs (or reads) is cumulative, since all ASVs assigned to the species level also have an assignment for higher taxonomic ranks.

4 | DISCUSSION

ASV distributions obtained from groundwater eDNA samples from 20 sites in North Eastern Switzerland revealed different organismal communities according to the land-use type surrounding the SCB or according to the type of aquifer sampled. Groundwater communities in agricultural versus forested areas were significantly different (Figure 6, Table 2), and further differentiated between unconsolidated and fissured aquifers. Signals of land-use and geology on the chemical and physical properties of groundwater are well-established (e.g. Lerner & Harris, 2009 and references therein), but very few studies assessed their impacts on macro-organisms (e.g. Korbel et al., 2022 for microbial communities). Here we demonstrate, using eDNA, a clear imprint of land-use and geology on macro- and micro-organismal diversity and community structure in groundwater in a well-defined area. These results are consistent with the findings of Korbel et al. (2013), who showed variations in both microbial functional diversity and stygofaunal assemblages according to different agricultural practices in Australia. Since water quality is strongly dependent on ecosystem functions provided by stygobionts, which could in turn be impacted by surface land-use, our study highlights the relevancy of considering biological parameters, alongside chemical and physical parameters, when evaluating the state and change of groundwater systems.

We identified groundwater organisms from different taxonomic groups, including emblematic species such as groundwater amphipods of the genus *Niphargus*. However, 85% of reads could not be assigned to any taxa, even with a loose identity threshold of 80%.

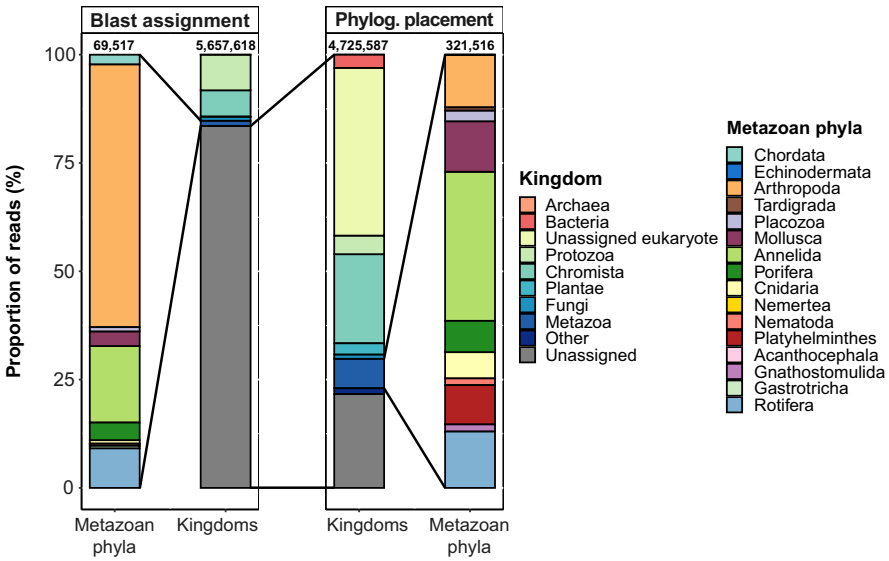

FIGURE 4 Distribution of read assignments at the kingdom level and for the different metazoan phyla observed, with our two assignment protocols: BLAST-based assignment on the left and phylogenetic placement of the unassigned portion of the Blast results on the right. The number on top of each bar represents the total number of reads for each category.

TABLE 1 List of the metazoan species identified after a blast-based taxonomic assignment of the sequences obtained with metabarcoding.

Phylum	Class	Order	Family	Assignment	GW affinity	% Ident	#Sites	#ASVs	#Reads	% Reads
Annelida	Clitellata	Crassiclitellata	Lumbricidae	*Aporrectodea longa*	Non-aquatic	100.0	1	1	31	0.001
Annelida	Clitellata	Enchytraeida	Enchytraeidae	*Cernosvitoviella aggtelekiensis*	Stygophile	94.6	1	1	10	<0.001
Annelida	Clitellata	Enchytraeida	Enchytraeidae	*Enchytraeus bulbosus*	Non-aquatic	99.0	1	1	705	0.012
Annelida	Clitellata	Enchytraeida	Enchytraeidae	*Fridericia tuberosa*	Non-aquatic	96.1	1	1	118	0.002
Annelida	Clitellata	Tubificida	Naididae	*Rhyacodrilus falciformis*	Stygophile	100.0	1	1	54	0.001
Arthropoda	Arachnida	Araneae	Linyphiidae	*Microneta viaria*	Non-aquatic	100.0	2	1	66	0.001
Arthropoda	Arachnida	Araneae	Theridiidae	*Paidiscura pallens*	Non-aquatic	100.0	1	1	447	0.008
Arthropoda	Arachnida	Araneae	Theridiidae	*Phylloneta impressa*	Non-aquatic	100.0	1	1	324	0.006
Arthropoda	Arachnida	Sarcoptiformes	Phenopelopidae	*Eupelops acromios*	Non-aquatic	93.1	1	1	41	0.001
Arthropoda	Diplopoda	Polydesmida	Polydesmidae	*Propolydesmus helveticus*	Non-aquatic	100.0	1	1	4	<0.001
Arthropoda	Diplopoda	Polyxenida	Polyxenidae	*Polyxenus lagurus*	Non-aquatic	100.0	1	1	21	<0.001
Arthropoda	Hexanauplia	Cyclopoida	Cyclopidae	*Acanthocyclops vernalis/robustus*	Stygophile	100.0	1	1	16,110	0.280
Arthropoda	Insecta	Diptera	Chironomidae	*Limnophyes habilis*	Stygoxene	100.0	1	1	42	0.001
Arthropoda	Insecta	Diptera	Syrphidae	*Eristalis pertinax*	Stygoxene	100.0	1	1	20	<0.001
Arthropoda	Insecta	Diptera	Tipulidae	*Tipula paludosa*	Stygoxene	100.0	1	1	11	<0.001
Arthropoda	Insecta	Hemiptera	Aphididae	*Rhopalosiphoninus latysiphon*	Non-aquatic	100.0	1	1	13	<0.001
Arthropoda	Insecta	Hymenoptera	Ichneumonidae	*Enclisis vindex*	Non-aquatic	98.5	1	1	93	0.002
Arthropoda	Malacostraca	Amphipoda	Niphargidae	*Niphargus auerbachi*	Stygobite	100.0	1	1	23	<0.001
Arthropoda	Malacostraca	Amphipoda	Niphargidae	*Niphargus fontanus*	Stygobite	100.0	2	2	412	0.007
Arthropoda	Malacostraca	Amphipoda	Niphargidae	*Niphargus puteanus*	Stygobite	99.5	1	1	254	0.004
Arthropoda	Malacostraca	Amphipoda	Niphargidae	*Niphargus thienemanni*	Stygobite	100.0	1	2	284	0.005
Arthropoda	Malacostraca	Amphipoda	Niphargidae	*Niphargus tonywhitteni*	Stygobite	100.0	3	1	76	0.001
Arthropoda	Malacostraca	Cumacea	Nannastacidae	*Campylaspis sulcata*	Stygoxene	92.7	1	2	108	0.002
Arthropoda	Protura	None	Fujientomidae	*Fujientomon dicestum*	Non-aquatic	100.0	1	1	9	<0.001
Chordata	Mammalia	Artiodactyla	Suidae	*Sus scrofa*	Non-aquatic	100.0	1	2	234	0.004
Chordata	Mammalia	Primates	Hominidae	*Homo sapiens*	Non-aquatic	100.0	2	2	478	0.008
Chordata	Mammalia	Rodentia	Muridae	*Mus musculus*	Non-aquatic	100.0	7	1	864	0.015
Mollusca	Gastropoda	Stylommatophora	Agriolimacidae	*Deroceras laeve*	Non-aquatic	99.5	1	1	240	0.004
Mollusca	Gastropoda	Stylommatophora	Boettgerillidae	*Boettgerilla pallens*	Non-aquatic	100.0	1	1	688	0.012

TABLE 1 (Continued)

Phylum	Class	Order	Family	Assignment	GW affinity	% Ident	#Sites	#ASVs	#Reads	% Reads
Nematoda	Chromadorea	Rhabditida	Alloionematidae	*Alloionema appendiculatum*	Non-aquatic	97.5	1	1	282	0.005
Rotifera	Eurotatoria	None	None	*Rotaria rotatoria*	Stygoxene	100.0	1	1	132	0.002
Rotifera	Eurotatoria	Philodinida	Philodinidae	*Macrotrachela quadricornifera*	Stygoxene	91.8	1	1	139	0.002
Rotifera	Eurotatoria	Philodinida	Philodinidae	*Philodina flaviceps*	Stygoxene	99.5	1	1	29	0.001

Note: The detailed taxonomy of each species is specified as well as their affinity for groundwater according to Gibert et al. (1994). Stygobites are specialized subterranean organisms, obligatory hypogean, stygophiles are organisms that can spend part of their life in subterranean environments, and stygoxenes are organisms with no affinity for groundwater. For each assignment, the identity between the sequence and the reference is given as well as the number of ASVs and reads attributed to the species, the number of sites where these ASVs were found, and the proportion of reads within the whole data set.

Thus, the observed patterns of diversity connected to land-use were largely driven by a large fraction of undescribed organismal diversity (or at least unrecorded in the databases). Gaps in reference databases are a general limitation for metabarcoding studies (Li et al., 2022; Weigand et al., 2019), and are especially pronounced in under-studied ecosystems such as groundwater or soil.

4.1 | Validation of our protocol

To validate the observed diversity patterns, we needed to ensure that our eDNA metabarcoding protocol is adequate for groundwater biodiversity assessments and not composed of technical artefacts. From the first uses of eDNA-based methods in surface waters, studies focused on comparing their results with traditional sampling approaches (Keck et al., 2022 and references therein). For subterranean environments, however, such a validation is more difficult as very few of them have already been investigated by a traditional approach (Mammola et al., 2019). It is thus imperative to get enough replication for an internal evaluation of the method. Previous studies using eDNA in subterranean habitats applied protocols derived from what is done for surface environments, with a volume of water filtered around 1–2 L per replicate, and either none or a low number of technical replicates (e.g. Korbel et al., 2017; Niemiller et al., 2018; West et al., 2020). From their results, however, it seems that these sampling efforts are insufficient when targeting subterranean taxa. As the amount of organisms living in groundwater is thought to be low (especially for multicellular eukaryotes; Deharveng et al., 2009), and because the water collected in SCB contains a low amount of sediments and other particulates to which DNA can bind itself (Kumar et al., 2022), we maximized the volume of water filtered (4 × 10 L) and the number of replicates at the amplification step (4 filters × 5 PCRs). The number of ASVs retrieved when pooling all replicates reached a plateau for all sampled sites, sometimes even before attaining 15 replicates (Figure 2b), demonstrating that our protocol was sufficient to retrieve a representative view of the amplified biodiversity. However, when looking into detail at the effect of the different type of replicates we can see that the number of PCR replicates falls slightly short to cover the whole richness present in each filter (Figure 2d), although the number of filters used was sufficient (Figure 2c). This result, combined with the fact that PCR replicates are more similar to each other than are filter replicates (Figure 5), suggest that the ASVs divergently amplified in the different PCR replicates are in low abundance. This could be explained by the persistence of PCR errors in our data set despite our correction step or by the presence of numerous rare species in our samples. The latter is congruent with the numerous low-abundance metazoan ASVs assigned to the species level (Table 1). To increase cost-efficiency, we thus advise to choose the number of replicates depending on the question at hand, favouring a high number of filter replicates when interested in broad diversity patterns, and increasing the number of PCR replicates

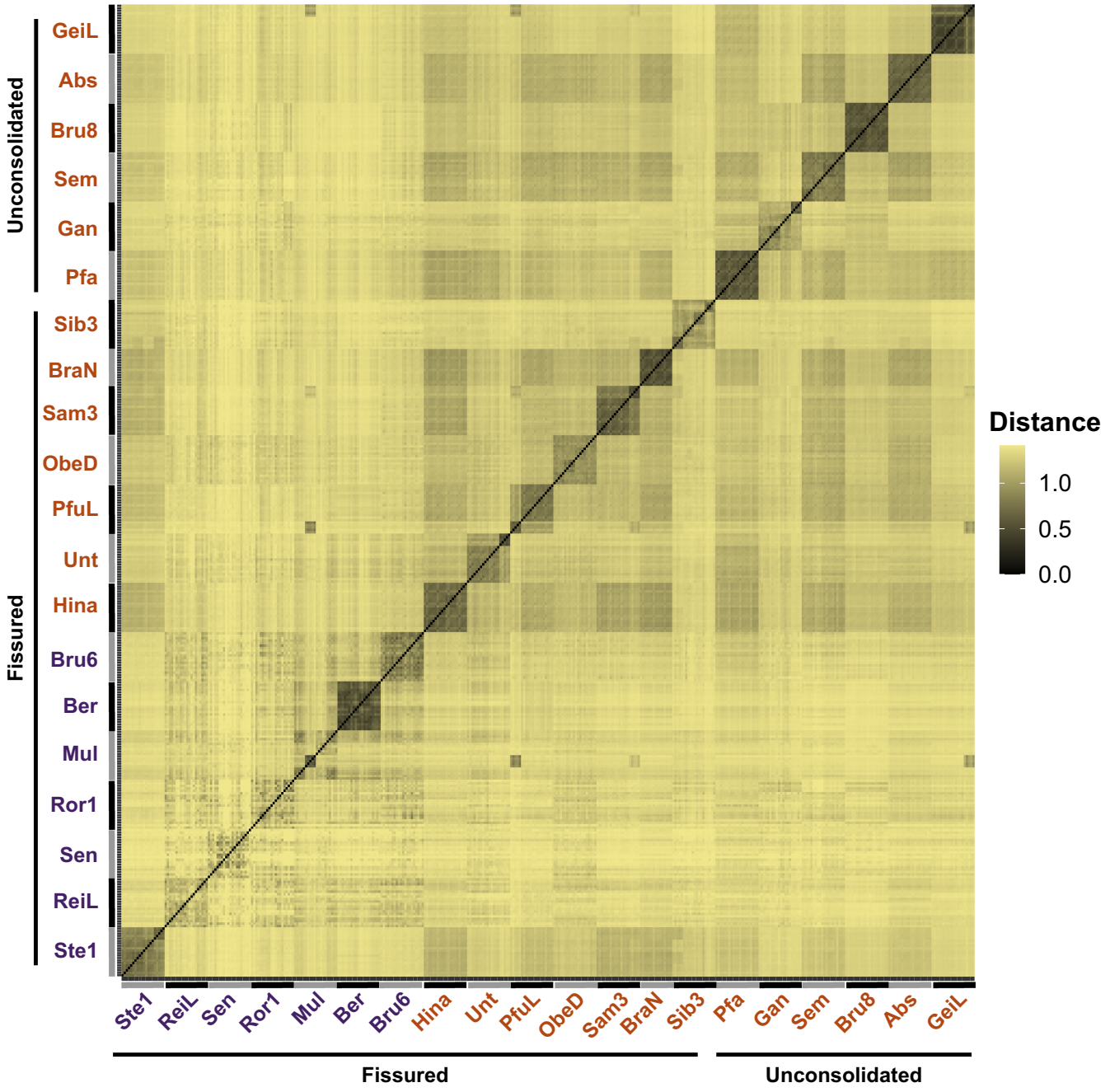

FIGURE 5 Heatmap of the Euclidean distance between the 400 replicates of our study, calculated on the ASV distribution after Hellinger transformation. All 20 replicates of a same site are side by side. The label colours indicate surrounding land-use at the sampling site, which was forest (orange) or agriculture (purple). The two types of aquifers are indicated (fissured or unconsolidated). See Figure 1 for site codes.

when focusing on species detection, especially metazoan species. Whatever the choice, we still suggest to keep at least three tagged PCR replicates, as they are particularly useful to remove potential PCR errors.

When considering sequencing coverage, one MiSeq run was sufficient to sequence the vast majority of DNA signals present in our library (Figure 2a). The high coverage variability between sites (min: 54,808 reads; max: 650,792 reads) reflects the amplification success, with the samples with the lowest number of reads being the ones for which no PCR product could be observed on gel pictures. This is probably representative of the scarcity of organisms in these sites, with a lower amount of DNA available and a lower diversity (Table S3). We believe that this is the most likely hypothesis (as opposed to a lack of amplification due to a technical artefact) because all replicates from a same site did consistently fail or poorly amplify, and because previous unpublished observations were not able to collect macroinvertebrates at those sites.

While all rarefaction curves reached a plateau rather quickly (Figure 2a), all identified metazoan species are associated with a very low number of reads (Table 1). This issue, stemming from

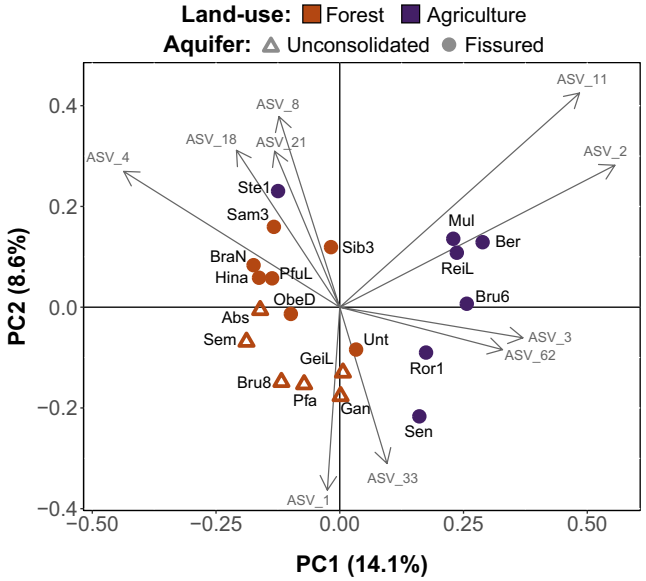

FIGURE 6 Principal component analysis (PCA) of the 20 studied sites based on ASV distribution data after Hellinger transformation. The colours indicate the type of land-use surrounding the sampling site: orange for forest and purple for agriculture. Filled circles and open triangles indicate that the water comes from fissured or unconsolidated aquifers, respectively. The ten ASVs contributing the most to the different axes are displayed with grey arrows. See Figure 1 for site codes.

TABLE 2 Results of the PERMANOVA performed on the Euclidean distance between sites calculated from the ASV distribution table (all replicates pooled per site) transformed through the 'Hellinger' method.

	Degrees of freedom	Sum of squares	*F*	R^2	*p*-Value
a					
Month	2	1.377	1.063	.104	.253
Aquifer	1	0.990	1.528	.075	.012
Land-use	1	1.168	1.804	.088	.002
Residuals	15	9.715	–	.733	–
b					
Month	2	1.377	1.063	.104	.252
Land-use	1	1.463	2.258	.110	<.001
Aquifer	1	0.696	1.074	.053	.263
Residuals	15	9.715	–	.733	–

Note: These were obtained after 9999 permutations. All terms were added sequentially and the results of either the impact of land-use after considering month and aquifer (a) or the impact of aquifer after considering month and land-use (b) are indicated.

the low concentration of metazoan DNA in water (as compared to micro-organisms), should be addressed at the amplification step, by using primers specifically targeting metazoan taxa. Our attempt to address this issue by using existing primers designed to target freshwater macroinvertebrates (Vamos et al., 2017) was not fully successful, even after modifying them to ensure the amplification of groundwater amphipod species from the genus *Niphargus*. The original primers are known to exhibit a low amplification of prokaryotic groups and to have a strong affinity for metazoans, but they can also amplify non-target taxa, especially from other kingdoms such as Chromista, Fungi or Protozoa (e.g. Hupało et al., 2022). As the modifications consisted of the addition of wobble bases, it was expected that our primers would amplify a higher number of taxa than the original version. The results from an in silico PCR, however, performed on all sequences available in the GenBank nt database, revealed that they should target mostly metazoan taxa (Figure S1). This is thus surprising that the most numerous groups identified in our data set are micro-organisms (Protozoa and Chromista; Figure 4). This could be explained either by the very high concentration of micro-eukaryotes in our samples, which would lead to their amplification despite a higher number of mismatches with the primers. Another likely explanation could be the lack of references in GenBank's nt database for the micro-eukaryotes identified in our study, thus biasing the results of the in silico PCR.

4.2 | eDNA metabarcoding is effective in identifying groundwater macro-organisms

Among the 4917 ASVs listed with our approach, 752 (15.3%) could be assigned to a taxon, across five eukaryote kingdoms and 11 metazoan phyla (Figure 4). However, only 89 (1.8%) could be assigned at the species level. The lack of resolution in our assignments is due to the low number of reference sequences available in public databases for groundwater taxa. For example, among the 827 species of molluscs, arthropods and annelids listed in the PASCALIS project report (Stoch et al., 2004), only 211 (25.5%) have at least one sequence available in GenBank, and 186 (22.5%) have been sequenced for COI. We could expect that the representation of these organisms would be higher in more curated databases such as BOLD (Ratnasingham & Hebert, 2007), and this is why we tried to perform the same assignment protocol using this database as well (Table S5). Contrary to the expectations, the use of BOLD resulted in the same amount of ASVs assigned (13% at 80% identity vs. 15% with GenBank, and 3% at 90% identity vs. 4% with GenBank). The combination of both databases would only increase the number of metazoan species identified by 13, none of which having an affinity for groundwater. Even though the amount of reads assigned to metazoans, and particularly to arthropods, is higher with BOLD (Figure S2), we can wonder if this is not an artefact due to the low representation of other groups in the database. In fact 85% of specimens present in BOLD are arthropods.

A low representation of subterranean organisms in public databases is reinforced as many species in subterranean environments are not yet discovered or described (Zagmajster et al., 2018). This issue is reported by other authors applying metabarcoding approaches to groundwater (Korbel et al., 2017; West et al., 2020). It is thus not surprising that our assignment level was so low and

that only 62 species could be identified. Additionally, the presence of particular groundwater species might have been missed due to unknown amplification biases.

About half of the species (33) recovered in our metabarcoding data set are metazoan species, from five different phyla (Table 1). Despite our use of a relatively loose threshold of 90% for species assignment, all species but three have already been recorded in Switzerland or in neighbouring countries, making their presence in our data set plausible. Only *Campylapsis sulcata*, a marine species, *Alloionema appendiculatum*, a nematode from North America, and *Fujientomon dicestum*, a Protura from China, are unlikely. For the two first, their lower identity with the references probably indicates that the DNA present in our samples belongs to a closely related species. For the latter, however, the identity of 100% would point towards the presence of an erroneous sequence in the database. Among the reliable assignments, five species are obligatory groundwater species (stygobites) and belong to the amphipod genus *Niphargus*, all already observed in Switzerland (Altermatt et al., 2019; Alther et al., 2021). Three other species are stygophiles, meaning that they can occasionally spend part of their life in groundwater. One of them is the copepod species complex *Acanthocyclops vernalis/robustus*, which is a very widespread taxon in the northern hemisphere and which has been observed both in surface and ground-waters (Galassi, 2001). *Cernosvitoviella aggtelekiensis*, an annelid species known to occasionally occur in groundwater, was recorded only one time in Switzerland under the synonym *C. goodhui* (Dumnicka et al., 2015). Our ASV, however, is only 94.6% similar to the reference sequence, and could belong to a closely related species, not listed in GenBank. The last stygophile species is *Rhyacodrilus falciformis*, another annelid taxon that was originally described from Switzerland (Bretscher, 1901). The remaining 25 species are stygoxenes or non-aquatic taxa, meaning that they have no affinity for groundwater or live in the soil or in surface habitats. The presence of their DNA in our samples could come from a contamination at the time of sampling as some of them lived in the SCB (e.g. spiders). Contamination during filtration or lab processing cannot be completely excluded (especially for human DNA) even if none of the controls exhibited any reads after correction. To our opinion, however, the most likely explanation for most of them is that these organisms (or part of them) were transported through the soil and into the groundwater during rain events. This theory is also mentioned by Lunghi et al. (2022) as an explanation for the presence of freshwater macroinvertebrates' DNA in cave soil. This is also congruent with the fact that other groundwater eDNA studies systematically found DNA from surface organisms in their samples (Oberprieler et al., 2021; West et al., 2020).

4.3 | Still sequencing in the dark

Despite having satisfactorily recovered several species from diverse habitats, the majority of our ASVs (96%) remained unassigned. Even when considering a loose assignment threshold of 80% identity, 85% of ASVs did not match with any references from GenBank. The presence of a high amount of unassigned sequences in eDNA metabarcoding data sets is a common issue, particularly with COI (Collins et al., 2019 and references therein). Their proportion in our study, however, is even higher than in other publications. For example, Macher et al. (2018), although using the same identity threshold as our study (80%) could assign 37% of their OTUs (vs. 15% of our ASVs) from river eDNA in New-Zealand. Similarly, in Couton et al. (2022), 24% of reads from eDNA collected in marinas could be assigned to a metazoan taxon, with an even more stringent identity threshold (92%), when the number of reads assigned to metazoans in our study is 1.4%. These comparisons suggest that a factor, specific for groundwater, is preventing many assignments.

Given the very high percentage of unassigned ASVs, one could wonder if these reads are actual DNA sequences present in our samples or if they could be technical artefacts. Metabarcoding is prone to several types of errors such as PCR errors, sequencing errors, or chimeras (Alberdi et al., 2018 and references therein). All these issues, however, arise stochastically during the library construction, are less abundant than the 'true' sequences from which they derive, and are expected to occur randomly across the samples/replicates. Contrastingly, the most abundant ASVs are unassigned in our data set (Table S4), the ASV distribution is very consistent between replicates (Figure 5), and there is a higher similarity within replicates of one sampled site than between sites. Another source of bias comes from the ability of universal primers to amplify parts of the genome that are not targeted. This is particularly relevant when focusing on mitochondrial markers, because of the presence of NUMTs (i.e. nuclear non-functional copies of mitochondrial genes; Bensasson et al., 2001). Here again, NUMTs are supposed to be less abundant than mitochondrial sequences because they are present in a lower copy number within the cell (Andújar et al., 2021), which does not fit the pattern of our data set. Moreover, NUMTs sequences usually contain non-synonymous mutations, which can produce stop-codons (Andújar et al., 2021). To control for such an artefact, we translated all sequences assigned to arthropods with the phylogenetic placement method (Table S6). Almost all (99.3%) of them did translate into protein sequences with only one of them including a stop-codon. Furthermore, all of these translated sequences exhibited a protein sequence highly similar to COI sequences, revealing that the presence of NUMTs is only minor in our data set. Together, this corroborates that the sequences produced with our protocol are actual true COI sequences from organisms not referenced in public databases.

The high number of unassigned sequences with COI is usually thought to be the result of convergence between the COI gene and its homologues in prokaryotes. For example, Siddall et al. (2009) demonstrated that marine gammaproteobacteria can be successfully amplified by the universal barcoding primers designed by Folmer et al. (1994), and supposed to target exclusively metazoans. The results of the phylogenetic placement approach on the unassigned portion of our data set (Figure 4; Table S7) shows that only a small proportion of reads were assigned to Bacteria (2.6%) or Archaea (<0.1%). Most of our data set seems to be composed of

micro-eukaryotes. The lack of reference sequences for these organisms is not surprising as many of them are difficult to identify morphologically, and because COI is usually not the marker of choice for these groups (Burki et al., 2021; Zimmermann et al., 2015). More surprising are the 6.8% of unassigned sequences attributed to metazoans. This result underlines the lack of knowledge on groundwater macro-organisms, and calls for a greater effort for describing new species but also for sequencing existing specimens, as advocated by Saccò et al. (2022). An online database dedicated to subterranean taxa, called Stygofauna Mundi, is in development (Martínez et al., 2018), which could stimulate the discovery of new species and could help future eDNA metabarcoding works.

4.4 | Diversity present in groundwater eDNA reflects land-use and geology

Given that taxonomic assignment is limiting the understanding of groundwater eDNA sequences, we used a taxonomy-free approach and looked at the diversity of ASVs within and between sites. The community composition in the sampled sites were clearly different depending on the land-use (i.e. if the SCB was situated in a forest or in an agricultural area) and on the aquifer type (Figure 6). This result was further confirmed by the PERMANOVA for land-use (Table 2), with a significant effect on ASV distributions. Moreover, the sites located in an agricultural area had a consistently lower ASV richness (Table S3). The only exception to this pattern is Ste1, a site located just downhill from a golf course. We classified it as 'agriculture' because we believed that the intensive care of the golf lawns would have a similar effect on soil and groundwater communities than crop growing. It seems, however, that this is not the case as the eDNA pattern is closer to forested sites.

The diversity patterns were strong, despite the relatively small fraction of the variance between sites explained (Table 2). Many parameters could explain the remaining 73.3% of unexplained variance, which were not addressed in our study. For example, it could be linked to several chemical factors, such as the presence of particular pollutants or the amount of dissolved oxygen, or to physical parameters, such as the temperature, the flow rate at the time of sampling, past drying events of the aquifer or the size of pore spaces in the sediment. For the latter, the classification of aquifer types used in this study was very coarse, and within each category, aquifers can have different geological compositions with various porosities. These micro-scale changes could play an important role in the availability of suitable habitats, especially for macro-organisms. Many studies demonstrated the impact of land-use and geology on chemical and physical properties of groundwater (Foster & Custodio, 2019; Lerner & Harris, 2009). Micro-organisms are also often targeted to show the biological impact of pollution or depletion (e.g. Korbel et al., 2022; Kwon et al., 2020). We show here that all these modifications in groundwater quality and quantity are also reflected in the eukaryotic community, which could serve as a proxy for a more integrative monitoring, as also suggested by Korbel et al. (2013).

Out of the 10 ASVs that contribute the most to the observed pattern, seven are unassigned (Figure 6; Table S4). The three others are associated with protozoan genera from the phylum Amoeboza (*Vexillifera*, *Vannella*, and *Korotnevella*), which are not known to be living in groundwater. As in any taxonomy-free approach, we cannot know if the patterns observed are attributed to stygobites, stygophiles or stygoxenes/non-aquatic taxa. If the proportion of each of these groups identified in Table 1 is representative of the whole data set, it is likely that most of our ASVs are actually associated with stygoxene and non-aquatic taxa, possibly washed in from surface/soil ecosystems. However, a tight coupling of above- and below-ground ecosystems is expected, and groundwater signals should integrate all these communities, as demonstrated by Lunghi et al. (2022). The differences observed related to land-use could be attributed to soil micro-organisms, but since all strata (i.e. surface, soil, and groundwater) are interconnected, if one is impacted, it is likely that the others will be too. Based on previous knowledge about subterranean organisms from caves (e.g. Fong, 2019; Simon et al., 2003), we can assume that the major food source of stygobionts is coming from the surface or the soil. The organic matter detected in our samples, although belonging to stygoxene/non-aquatic organisms, could thus be utilized by groundwater species, and be mechanistically coupled to groundwater communities.

In conclusion, we demonstrated that eDNA metabarcoding could be a valuable tool in the study of underground communities. Although our lack of knowledge on the organisms living in groundwater can dramatically impair taxonomic assignments of metabarcoding sequences, the use of a taxonomy-free approach allowed us to reveal an abundant unknown diversity. The comparison of the observed genetic assemblages revealed a difference between communities associated with forested or agricultural areas as well as between the different aquifers sampled. This result, although expected since many chemical and physical factors are already known to differ depending on land-use and geology, is demonstrating the necessity to consider biodiversity in groundwater management plans. Since most of the taxa identified in our data set were surface or soil organisms, we believe that eDNA-based monitoring could be particularly interesting in that regard as it would allow an integrative view of several interacting communities. Of course, more work is needed to evaluate the capacity of groundwater eDNA to adequately depict each community but we strongly believe that it should be included in surveillance programmes to investigate the impact of land-use or climate change on the aquifers.

AUTHOR CONTRIBUTIONS

Florian Altermatt, Roman Alther, and Marjorie Couton designed the experiment, Angela Studer, Roman Alther, and Marjorie Couton collected the samples, Samuel Hürlemann and Marjorie Couton processed the samples, Marjorie Couton analysed the data, Florian Altermatt and Marjorie Couton led the writing of the manuscript. All authors contributed critically to the drafts and gave final approval for publication.

ACKNOWLEDGEMENTS

We thank all water providers, Mara Knüsel and Anouk Petitpierre for their help with the sampling. Data produced and analysed in this paper were generated in collaboration with the Genetic Diversity Centre (GDC), ETH Zurich. We thank three anonymous reviewers for their constructive comments on a previous version of the manuscript. Funding for this project was provided by the Swiss National Science Foundation (grant nr. 31003A_173074), the Swiss Federal Office for the Environment (FOEN/BAFU), the Fachverband für Gas, Wärme und Wasser (SVGW/SSIGE) and the University of Zurich Research Priority Programme in Global Change and Biodiversity (URPP GCB). Open access funding provided by ETH-Bereich Forschungsanstalten.

BENEFIT-SHARING STATEMENT

We consulted with local water providers for groundwater biodiversity assessments that granted us access to their spring catchment boxes. The contributions of all individuals to the research are described in the METHODS and ACKNOWLEDGEMENTS. Benefits from this research accrue from the sharing of our data and results on public databases as described above.

DATA AVAILABILITY STATEMENT

Raw sequence reads are deposited on NCBI, in the SRA: BioProject PRJNA905821. The metabarcoding and analyses scripts used in this study are available at https://github.com/joarwrie/NiphToess. Sampling metadata are provided in Table S1 but due to the sensitivity of drinking water well locations, exact coordinates can only be given upon request to the authors.

ORCID

Marjorie Couton https://orcid.org/0000-0001-9880-8646
Roman Alther https://orcid.org/0000-0001-7582-3966
Florian Altermatt https://orcid.org/0000-0002-4831-6958

REFERENCES

Alberdi, A., Aizpurua, O., Gilbert, M. T. P., & Bohmann, K. (2018). Scrutinizing key steps for reliable metabarcoding of environmental samples. *Methods in Ecology and Evolution*, *9*(1), 134–147. https://doi.org/10.1111/2041-210x.12849

Altermatt, F., Atlher, R., Fišer, C., & Švara, V. (2019). *Amphipoda (Fauna Helvetica)* (Vol. *32*). University of Zurich. https://www.zora.uzh.ch/id/eprint/174616/

Alther, R., Bongni, N., Borko, Š., Fišer, C., & Altermatt, F. (2021). Citizen science approach reveals groundwater fauna in Switzerland and a new species of *Niphargus* (Amphipoda, Niphargidae). *Subterranean Biology*, *39*, 1–31. https://doi.org/10.3897/subtbiol.39.66755

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, *215*(3), 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2

Andújar, C., Creedy, T. J., Arribas, P., López, H., Salces-Castellano, A., Pérez-Delgado, A. J., Vogler, A. P., & Emerson, B. C. (2021). Validated removal of nuclear pseudogenes and sequencing artefacts from mitochondrial metabarcode data. *Molecular Ecology Resources*, *21*(6), 1772–1787. https://doi.org/10.1111/1755-0998.13337

Bensasson, D., Zhang, D.-X., Hartl, D. L., & Hewitt, G. M. (2001). Mitochondrial pseudogenes: Evolution's misplaced witnesses. *Trends in Ecology & Evolution*, *16*(6), 314–321. https://doi.org/10.1016/S0169-5347(01)02151-6

Benson, D. A., Cavanaugh, M., Clark, K., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., & Sayers, E. W. (2013). GenBank. *Nucleic Acids Research*, *41*(Database issue), D36–D42. https://doi.org/10.1093/nar/gks1195

Borcard, D., Gillet, F., & Legendre, P. (2018). *Numerical ecology with R* (2nd ed.). Springer. https://doi.org/10.1007/978-3-319-71404-2

Boulton, A. J., Fenwick, G. D., Hancock, P. J., & Harvey, M. S. (2008). Biodiversity, functional roles and ecosystem services of groundwater invertebrates. *Invertebrate Systematics*, *22*(2), 103–116. https://doi.org/10.1071/IS07024

Boyer, F., Mercier, C., Bonin, A., Le Bras, Y., Taberlet, P., & Coissac, E. (2016). OBITOOLS: A UNIX-inspired software package for DNA metabarcoding. *Molecular Ecology Resources*, *16*(1), 176–182. https://doi.org/10.1111/1755-0998.12428

Brantschen, J., Blackman, R. C., Walser, J.-C., & Altermatt, F. (2021). Environmental DNA gives comparable results to morphology-based indices of macroinvertebrates in a large-scale ecological assessment. *PLoS One*, *16*(9), e0257510. https://doi.org/10.1371/journal.pone.0257510

Brantschen, J., Pellissier, L., Walser, J.-C., & Altermatt, F. (2022). Evaluation of primer pairs for eDNA-based assessment of Ephemeroptera, Plecoptera, and Trichoptera across a biogeographically diverse region. *Environmental DNA*, *4*(6), 1356–1368. https://doi.org/10.1002/edn3.342

Bretscher, K. (1901). Beobachtungen uber Oligochæten der Schweiz. *Revue Suisse de Zoologie*, *9*, 189–223.

Burki, F., Sandin, M. M., & Jamy, M. (2021). Diversity and ecology of protists revealed by metabarcoding. *Current Biology*, *31*(19), R1267–R1280. https://doi.org/10.1016/j.cub.2021.07.066

Burri, N. M., Weatherl, R., Moeck, C., & Schirmer, M. (2019). A review of threats to groundwater quality in the anthropocene. *Science of the Total Environment*, *684*, 136–154. https://doi.org/10.1016/j.scitotenv.2019.05.236

Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., & Holmes, S. P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. *Nature Methods*, *13*, 581–583. https://doi.org/10.1038/nmeth.3869

Collins, R. A., Bakker, J., Wangensteen, O. S., Soto, A. Z., Corrigan, L., Sims, D. W., Genner, M. J., & Mariani, S. (2019). Non-specific amplification compromises environmental DNA metabarcoding with COI. *Methods in Ecology and Evolution*, *10*(11), 1985–2001. https://doi.org/10.1111/2041-210x.13276

Couton, M., Lévêque, L., Daguin-Thiébaut, C., Comtet, T., & Viard, F. (2022). Water eDNA metabarcoding is effective in detecting non-native species in marinas, but detection errors still hinder its use for passive monitoring. *Biofouling*, *38*(4), 367–383. https://doi.org/10.1080/08927014.2022.2075739

Deharveng, L., Stoch, F., Gibert, J., Bedos, A., Galassi, D., Zagmajster, M., Brancelj, A., Camacho, A., Fiers, F., Martin, P., Giani, N., Magniez, G., & Marmonier, P. (2009). Groundwater biodiversity in Europe. *Freshwater Biology*, *54*(4), 709–726. https://doi.org/10.1111/j.1365-2427.2008.01972.x

Deiner, K., Bik, H. M., Mächler, E., Seymour, M., Lacoursière-Roussel, A., Altermatt, F., Creer, S., Bista, I., Lodge, D. M., de Vere, N., Pfrender, M. E., & Bernatchez, L. (2017). Environmental DNA metabarcoding: Transforming how we survey animal and plant communities. *Molecular Ecology*, *26*(21), 5872–5895. https://doi.org/10.1111/mec.14350

Deiner, K., Walser, J.-C., Mächler, E., & Altermatt, F. (2015). Choice of capture and extraction methods affect detection of freshwater biodiversity from environmental DNA. *Biological Conservation*, *183*, 53–63. https://doi.org/10.1016/j.biocon.2014.11.018

Dumnicka, E., Steingruber, S., Colombo, L., Zaupa, S., & Boggero, A. (2015). Oligochaete assemblages of Swiss Alpine lakes. *Italian Journal of Zoology*, *82*(1), 112–123. https://doi.org/10.1080/11250003.2014.965230

Elbrecht, V., & Leese, F. (2017). Validation and development of COI metabarcoding primers for freshwater macroinvertebrate bioassessment. *Frontiers in Environmental Science*, *5*, 11. https://doi.org/10.3389/fenvs.2017.00011

Ficetola, G. F., Canedoli, C., & Stoch, F. (2019). The Racovitzan impediment and the hidden biodiversity of unexplored environments. *Conservation Biology*, *33*(1), 214–216. https://doi.org/10.1111/cobi.13179

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.

Fong, D. W. (2019). Chapter 51 – Food sources. In W. B. White, D. C. Culver, & T. Pipan (Eds.), *Encyclopedia of caves* (3rd ed., pp. 429–434). Academic Press. https://doi.org/10.1016/B978-0-12-814124-3.00051-0

Foster, S., & Custodio, E. (2019). Groundwater resources and intensive agriculture in Europe – Can regulatory agencies cope with the threat to sustainability? *Water Resources Management*, *33*(6), 2139–2151. https://doi.org/10.1007/s11269-019-02235-6

Galassi, D. M. P. (2001). Groundwater copepods: Diversity patterns over ecological and evolutionary scales. *Hydrobiologia*, *453*(1), 227–253. https://doi.org/10.1023/A:1013100924948

Gibert, J., Danielopol, D. L., & Stanford, J. A. (1994). *Groundwater ecology (Aquatic Biology)* (Vol. *1*). Academic Press.

Griebler, C., & Avramov, M. (2015). Groundwater ecosystem services: A review. *Freshwater Science*, *34*(1), 355–367. https://doi.org/10.1086/679903

Herrmann, M., Geesink, P., Yan, L., Lehmann, R., Totsche, K. U., & Küsel, K. (2020). Complex food webs coincide with high genetic potential for chemolithoautotrophy in fractured bedrock groundwater. *Water Research*, *170*, 115306. https://doi.org/10.1016/j.watres.2019.115306

Hupało, K., Schmidt, S., Macher, T.-H., Weiss, M., & Leese, F. (2022). Fresh insights into Mediterranean biodiversity: Environmental DNA reveals spatio-temporal patterns of stream invertebrate communities on Sicily. *Hydrobiologia*, *849*(1), 155–173. https://doi.org/10.1007/s10750-021-04718-3

Keck, F., Blackman, R. C., Bossart, R., Brantschen, J., Couton, M., Hürlemann, S., Kirschner, D., Locher, N., Zhang, H., & Altermatt, F. (2022). Meta-analysis shows both congruence and complementarity of DNA and eDNA metabarcoding to traditional methods for biological community assessment. *Molecular Ecology*, *31*(6), 1820–1835. https://doi.org/10.1111/mec.16364

Koch, F., Menberg, K., Schweikert, S., Spengler, C., Hahn, H. J., & Blum, P. (2021). Groundwater fauna in an urban area – Natural or affected? *Hydrology and Earth System Sciences*, *25*(6), 3053–3070. https://doi.org/10.5194/hess-25-3053-2021

Korbel, K. L., Chariton, A., Stephenson, S., Greenfield, P., & Hose, G. C. (2017). Wells provide a distorted view of life in the aquifer: Implications for sampling, monitoring and assessment of groundwater ecosystems. *Scientific Reports*, *7*(1), 40702. https://doi.org/10.1038/srep40702

Korbel, K. L., Greenfield, P., & Hose, G. C. (2022). Agricultural practices linked to shifts in groundwater microbial structure and denitrifying bacteria. *Science of the Total Environment*, *807*, 150870. https://doi.org/10.1016/j.scitotenv.2021.150870

Korbel, K. L., Hancock, P. J., Serov, P., Lim, R. P., & Hose, G. C. (2013). Groundwater ecosystems vary with land use across a mixed agricultural landscape. *Journal of Environmental Quality*, *42*(2), 380–390. https://doi.org/10.2134/jeq2012.0018

Kumar, G., Farrell, E., Reaume, A. M., Eble, J. A., & Gaither, M. R. (2022). One size does not fit all: Tuning eDNA protocols for high- and low-turbidity water sampling. *Environmental DNA*, *4*(1), 167–180. https://doi.org/10.1002/edn3.235

Kwon, H.-I., Koh, D.-C., Jung, Y.-Y., Kim, D.-H., & Ha, K. (2020). Evaluating the impacts of intense seasonal groundwater pumping on stream–aquifer interactions in agricultural riparian zones using a multi-parameter approach. *Journal of Hydrology*, *584*, 124683. https://doi.org/10.1016/j.jhydrol.2020.124683

Lefébure, T., Douady, C. J., Gouy, M., Trontelj, P., Briolay, J., & Gibert, J. (2006). Phylogeography of a subterranean amphipod reveals cryptic diversity and dynamic evolution in extreme environments. *Molecular Ecology*, *15*(7), 1797–1806. https://doi.org/10.1111/j.1365-294X.2006.02888.x

Legendre, P., & Gallagher, E. D. (2001). Ecologically meaningful transformations for ordination of species data. *Oecologia*, *129*(2), 271–280. https://doi.org/10.1007/s004420100716

Leray, M., Yang, J. Y., Meyer, C. P., Mills, S. C., Agudelo, N., Ranwez, V., Boehm, J. T., & Machida, R. J. (2013). A new versatile primer set targeting a short fragment of the mitochondrial COI region for metabarcoding metazoan diversity: Application for characterizing coral reef fish gut contents. *Frontiers in Zoology*, *10*(1), 34. https://doi.org/10.1186/1742-9994-10-34

Lerner, D. N., & Harris, B. (2009). The relationship between land use and groundwater resources and quality. *Land Use Policy*, *26*, S265–S273. https://doi.org/10.1016/j.landusepol.2009.09.005

Li, F., Zhang, Y., Altermatt, F., Zhang, X., Cai, Y., & Yang, Z. (2022). Gap analysis for DNA-based biomonitoring of aquatic ecosystems in China. *Ecological Indicators*, *137*, 108732. https://doi.org/10.1016/j.ecolind.2022.108732

Lunghi, E., Valle, B., Guerrieri, A., Bonin, A., Cianferoni, F., Manenti, R., & Ficetola, G. F. (2022). Environmental DNA of insects and springtails from caves reveals complex processes of eDNA transfer in soils. *Science of the Total Environment*, *826*, 154022. https://doi.org/10.1016/j.scitotenv.2022.154022

Macher, J.-N., Vivancos, A., Piggott, J. J., Centeno, F. C., Matthaei, C. D., & Leese, F. (2018). Comparison of environmental DNA and bulk-sample metabarcoding using highly degenerate cytochrome c oxidase I primers. *Molecular Ecology Resources*, *18*(6), 1456–1468. https://doi.org/10.1111/1755-0998.12940

Mächler, E., Deiner, K., Steinmann, P., & Altermatt, F. (2014). Utility of environmental DNA for monitoring rare and indicator macroinvertebrate species. *Freshwater Science*, *33*(4), 1174–1183. https://doi.org/10.1086/678128

Mammola, S., Cardoso, P., Culver, D. C., Deharveng, L., Ferreira, R. L., Fišer, C., Galassi, D. M. P., Griebler, C., Halse, S., Humphreys, W. F., Isaia, M., Malard, F., Martinez, A., Moldovan, O. T., Niemiller, M. L., Pavlek, M., Reboleira, A. S. P. S., Souza-Silva, M., Teeling, E. C., … Zagmajster, M. (2019). Scientists' warning on the conservation of subterranean ecosystems. *Bioscience*, *69*(8), 641–650. https://doi.org/10.1093/biosci/biz064

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet Journal*, *17*(1), 10–12. https://doi.org/10.14806/ej.17.1.200

Martínez, A., Anicic, N., Calvaruso, S., Sánchez, N., Puppieni, L., Sforzi, T., Zaupa, S., Alvarez, F., Brankovits, D., Gąsiorowski, L., Gerovasileiou, V., Gonzalez, B., Humphreys, W., Iliffe, T., Worsaae, K., Bailly, N., & Fontaneto, D. (2018). A new insight into the Stygofauna Mundi: Assembling a global dataset for aquatic fauna in subterranean environments [Conference presentation abstract]. In *24th International Conference of Subterranean Biology*. University of Aveiro. https://doi.org/10.3897/aca.1.e29514

Morse, K. V., Richardson, D. R., Brown, T. L., Vangundy, R. D., & Cahoon, A. B. (2021). Longitudinal metabarcode analysis of karst bacterioplankton microbiomes provide evidence of epikarst to cave transport and community succession. *PeerJ*, *9*, e10757. https://doi.org/10.7717/peerj.10757

Mugnai, F., Costantini, F., Chenuil, A., Leduc, M., Gutiérrez Ortega, J. M., & Meglécz, E. (2023). Be positive: Customized reference databases and new, local barcodes balance false taxonomic assignments in metabarcoding studies. *PeerJ*, *11*, e14616. https://doi.org/10.7717/peerj.14616

Niemiller, M. L., Porter, M. L., Keany, J., Gilbert, H., Fong, D. W., Culver, D. C., Hobson, C. S., Kendall, K. D., Davis, M. A., & Taylor, S. J. (2018). Evaluation of eDNA for groundwater invertebrate detection and monitoring: A case study with endangered *Stygobromus* (Amphipoda: Crangonyctidae). *Conservation Genetics Resources*, *10*(2), 247–257. https://doi.org/10.1007/s12686-017-0785-2

Oberprieler, S., Rees, G., Nielsen, D., Shackleton, M., Watson, G., Chandler, L., & Davis, J. (2021). Connectivity, not short-range endemism, characterises the groundwater biota of a northern Australian karst system. *Science of the Total Environment*, *796*, 148955. https://doi.org/10.1016/j.scitotenv.2021.148955

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P. R., O'Hara, R. B., Simpson, G. L., Solymos, P., Stevens, M. H. H., Szoecs, E., & Wagner, H. (2018). *vegan: Community ecology package (version 2.5-2)*. https://CRAN.R-project.org/package=vegan

Pawlowski, J., Apothéloz-Perret-Gentil, L., & Altermatt, F. (2020). Environmental DNA: What's behind the term? Clarifying the terminology and recommendations for its future use in biomonitoring. *Molecular Ecology*, *29*(22), 4258–4264. https://doi.org/10.1111/mec.15643

Pfleger, M. O., Rider, S. J., Johnston, C. E., & Janosik, A. M. (2016). Saving the doomed: Using eDNA to aid in detection of rare sturgeon for conservation (Acipenseridae). *Global Ecology and Conservation*, *8*, 99–107. https://doi.org/10.1016/j.gecco.2016.08.008

Ratnasingham, S., & Hebert, P. D. N. (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

Saccò, M., Guzik, M. T., van der Heyde, M., Nevill, P., Cooper, S. J. B., Austin, A. D., Coates, P. J., Allentoft, M. E., & White, N. E. (2022). eDNA in subterranean ecosystems: Applications, technical aspects, and future prospects. *Science of the Total Environment*, *820*, 153223. https://doi.org/10.1016/j.scitotenv.2022.153223

Siddall, M. E., Fontanella, F. M., Watson, S. C., Kvist, S., & Erséus, C. (2009). Barcoding bamboozled by bacteria: Convergence to metazoan mitochondrial primer targets by marine microbes. *Systematic Biology*, *58*(4), 445–451. https://doi.org/10.1093/sysbio/syp033

Simon, K. S., Benfield, E. F., & Macko, S. A. (2003). Food web structure and the role of epilithic biofilms in cave streams. *Ecology*, *84*(9), 2395–2406. https://doi.org/10.1890/02-334

Smith, M., Cross, K., Paden, M., & Laban, P. (2016). *Spring–managing groundwater sustainably*. IUCN.

Somers, L. D., McKenzie, J. M., Mark, B. G., Lagos, P., Ng, G.-H. C., Wickert, A. D., Yarleque, C., Baraër, M., & Silva, Y. (2019). Groundwater buffers decreasing glacier melt in an andean watershed—But not forever. *Geophysical Research Letters*, *46*(22), 13016–13026. https://doi.org/10.1029/2019GL084730

Stoch, F., Malard, F., Castellarini, F., & Dole-Olivier, M.-J. (2004). *Statistical analyses and identification of indicators* (Vol. 7, pp. 1–156). PASCALIS D8 Deliverable for Workpackage.

Studer, A. (2022). *Diversity and distribution of groundwater amphipods in the catchment area of the river Töss* [unpublished master's thesis]. University of Zürich.

Sutherland, W. J., Butchart, S. H. M., Connor, B., Culshaw, C., Dicks, L. V., Dinsdale, J., Doran, H., Entwistle, A. C., Fleishman, E., Gibbons, D. W., Jiang, Z., Keim, B., Roux, X. L., Lickorish, F. A., Markillie, P., Monk, K. A., Mortimer, D., Pearce-Higgins, J. W., Peck, L. S., … Gleave, R. A. (2018). A 2018 horizon scan of emerging issues for global conservation and biological diversity. *Trends in Ecology & Evolution*, *33*(1), 47–58. https://doi.org/10.1016/j.tree.2017.11.006

Taberlet, P., Bonin, A., Zinger, L., & Coissac, E. (2018). *Environmental DNA: For biodiversity research and monitoring*. Oxford University Press. https://doi.org/10.1093/oso/9780198767220.001.0001

Taberlet, P., Coissac, E., Hajibabaei, M., & Rieseberg, L. H. (2012). Environmental DNA. *Molecular Ecology*, *21*(8), 1789–1793. https://doi.org/10.1111/j.1365-294X.2012.05542.x

Tang, Y., Wu, Y., Liu, K., Li, J., Li, H., Wang, Q., Yu, J., & Xu, P. (2019). Investigating the distribution of the Yangtze finless porpoise in the Yangtze River using environmental DNA. *PLoS One*, *14*(8), e0221120. https://doi.org/10.1371/journal.pone.0221120

Vamos, E. E., Elbrecht, V., & Leese, F. (2017). Short COI markers for freshwater macroinvertebrate metabarcoding. *Metabarcoding and Metagenomics*, *1*, e14625. https://doi.org/10.3897/mbmg.1.14625

Voisin, J., Cournoyer, B., Marjolet, L., Vienney, A., & Mermillod-Blondin, F. (2020). Ecological assessment of groundwater ecosystems disturbed by recharge systems using organic matter quality, biofilm characteristics, and bacterial diversity. *Environmental Science and Pollution Research*, *27*(3), 3295–3308. https://doi.org/10.1007/s11356-019-06971-5

Weigand, H., Beermann, A. J., Čiampor, F., Costa, F. O., Csabai, Z., Duarte, S., Geiger, M. F., Grabowski, M., Rimet, F., Rulik, B., Strand, M., Szucsich, N., Weigand, A. M., Willassen, E., Wyler, S. A., Bouchez, A., Borja, A., Čiamporová-Zaťovičová, Z., Ferreira, S., … Ekrem, T. (2019). DNA barcode reference libraries for the monitoring of aquatic biota in Europe: Gap-analysis and recommendations for future work. *Science of the Total Environment*, *678*, 499–524. https://doi.org/10.1016/j.scitotenv.2019.04.247

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. *Scientific Reports*, *10*(1), 21479. https://doi.org/10.1038/s41598-020-78525-6

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag. http://ggplot2.org

Zafeiropoulos, H., Gargan, L., Hintikka, S., Pavloudi, C., & Carlsson, J. (2021). The dark mAtteR iNvestigator (DARN) tool: Getting to know the known unknowns in COI amplicon data. *Metabarcoding and Metagenomics*, *5*, e69657. https://doi.org/10.3897/mbmg.5.69657

Zagmajster, M., Malard, F., Eme, D., & Culver, D. C. (2018). Subterranean biodiversity patterns from global to regional scales. In O. T. Moldovan, Ľ. Kováč, & S. Halse (Eds.), *Cave ecology* (pp. 195–227). Springer International Publishing. https://doi.org/10.1007/978-3-319-98852-8_9

Zimmermann, J., Glöckner, G., Jahn, R., Enke, N., & Gemeinholzer, B. (2015). Metabarcoding vs. morphological identification to assess diatom diversity in environmental studies. *Molecular Ecology Resources*, *15*(3), 526–542. https://doi.org/10.1111/1755-0998.12336

SUPPORTING INFORMATION

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

How to cite this article: Couton, M., Hürlemann, S., Studer, A., Alther, R., & Altermatt, F. (2023). Groundwater environmental DNA metabarcoding reveals hidden diversity and reflects land-use and geology. *Molecular Ecology*, *32*, 3497–3512. https://doi.org/10.1111/mec.16955