Alma Mater Studiorum – Università di Bologna

DOTTORATO DI RICERCA IN

SCIENZE DELLA TERRA, DELLA VITA E DELL'AMBIENTE

Ciclo 34

Settore Concorsuale: 05/B1 - ZOOLOGIA E ANTROPOLOGIA

Settore Scientifico Disciplinare: BIO/05 - ZOOLOGIA

The environmental DNA in the risk assessment and decision making processes for the invasive species management in agri-food sector, hydraulic security and biodiversity conservation

Presentata da:
Alex Cussigh

Coordinatore Dottorato:
Maria Giovanna Belcastro

Supervisore:
Barbara Mantovani

Co-Supervisori:
Giovanni Rossi e Federico Plazzi

Esame finale anno 2022

ABSTRACT

The use of environmental DNA (eDNA) analysis as a monitoring tool is becoming more and more widespread. The eDNA metabarcoding methods allow rapid community assessments of different target taxa.

This work is focused on the validation of the environmental DNA metabarcoding protocol for biodiversity assessment of freshwater habitats.

Scolo Dosolo was chosen as study area and three sampling points were defined for traditional and eDNA analyses. The gutter is a 205 m long anthropic canal located in Sala Bolognese (Bologna, Italy).

Fish community and freshwater invertebrate metazoans were the target groups for the analysis.

After a preliminary study in summer 2019, 2020 was devoted to the sampling campaign with winter (January), spring (May), summer (July) and autumn (October) surveys.

Alongside with the water samplings for the eDNA study, also traditional fish surveys using the electrofishing technique were performed to assess fish community composition; census on invertebrates was performed using an entomological net and a surber sampler.

After in silico analysis, the MiFish primer set amplifying a fragment of the *12s* rRNA gene was selected for bony fishes. For invertebrates the FWHF2 + FWHR2N primer combination, that amplifies a region of the mitochondrial *coi* gene, was chosen.

Raw reads were analyzed through a bioinformatic pipeline based on OBITools metabarcoding programs package and QIIME2. The OBITools pipeline retrieved seven fish taxa and 54 invertebrate taxa belonging to six different phyla, while QIIME2 recovered eight fish taxa and 45 invertebrate taxa belonging to the same six phyla as the OBITools pipeline.

The metabarcoding results were then compared with the traditional surveys data and bibliographic records.

Overall, the validated protocol provides a reliable picture of the biodiversity of the study area and an efficient support to the traditional methods.

SUMMARY

1 AIM OF THE THESIS

My PhD project was financed thanks to the 2014 - 2020 Regional Operational Programme (ROP) under the European Social Fund (ESF). The goal of the proposed project selected was to define and validate an environmental DNA metabarcoding protocol using freshwater samples in order to assess the biodiversity of inland waters in Emila-Romagna region. Once validated, the protocol could be also used to detect invasive species, even the ones that can compromise hydraulic security.

The study started with the identification of a 205 meters long stretch of the man-made canal Scolo Dosolo (Sala Bolognese, BO), belonging to the Special Protection Area (SPA) Cassa di espansione Dosolo, a Natura 2000 protected area (Habitat Directive; IT4050030), as study area. The choice was made since the site belongs to a protected area characterized by an interesting biodiversity component isolated from the surroundings.

Three sampling sites were defined along the canal and both traditional survey and eDNA methods were performed. The target taxa for this study were freshwater bony fishes and invertebrates.

During the summer 2019, a preliminary survey was carried out to test the efficiency of the different protocols and, later, the results were also used as a comparison with the data from the 2020 main campaign.

Even if the surveys have been complicated due to the covid outbreak, during the 2020 four seasonal campaigns were conducted. As stated before, in addition to the water samples for the eDNA metabarcoding analysis, traditional surveys were performed as well so as to have a comparison with the molecular data and to check for false positive/negative. In particular, three different sessions (summer 2019, summer 2020, autumn 2020) of electrofishing were performed by the Hydrosynergy co-op. staff, which also proceeded with the species identification. During the summer 2020 campaign, also a census on invertebrates was performed using entomological net and surber sampler; collected invertebrates were identified through morphological characters supported by standard barcoding of the cytochrome c oxidase I subunit gene for a

supplementary validation of the identification. Data from metabarcoding were compared with data from traditional surveys and bibliographic records.

The reads obtained from the sequencing of the MiFish primer set (Miya et al., 2015) amplifying a fragment of the *12s* rRNA gene, and the FWHF2 + FWHR2N primer combination (Vamos et al., 2017) for invertebrates, that amplifies a region of the mitochondrial *coi* gene, were analyzed.

Two pipelines, one based on the OBITools software package and the other one based on QIIME2, were used to check the ability of the different pipelines in retrieving biodiversity.

In general, the validated eDNA metabarcoding protocol provides a reliable picture of the biodiversity of the study area and an efficient support to the traditional methods.

Thanks to the collaboration with Hydrosynergy co-op. the protocol defined in Scolo Dosolo was also applied for the census and monitoring analyses of fish community and the elaboration of an updated management program in Natura 2000 sites of Laguna di Marano e Grado (SIC/ZPS IT3320037) and Valle Cavanata e Banco Mula di Muggia (SIC/ZPS IT3330006) in Grado (UD) and Marano Lagunare (UD).

In particular, 17 samples were collected during four different seasonal samplings and then analyzed using the OBITools and QIIME2 bioinformatic pipeline. These samples were included in the library preparation and sequencing of the MiFish reads of the 2020 campaign in Scolo Dosolo.

The results of these further analyses are not presented in this context, but they confirm the validation of the protocol outlined in Scolo Dosolo

2 INTRODUCTION

2.1 The eDNA metabarcoding

Environmental DNA (eDNA) designates the DNA which can be found and therefore extracted from environmental samples such as water, soil, air, snow, etc. (Taberlet et al., 2012). Total eDNA consists in both DNA shred from living organisms and extracellular DNA derived from cell death and subsequent cell degradation (Pietramellara et al., 2009).

Environmental DNA can be used as tool to detect a single species or multiple species, both approaches mainly based on Polymerase Chain Reaction (PCR; Taberlet et al., 2012). If the purpose is the detection of a single species, the best method is given by a quantitative PCR (qPCR). Otherwise, an approach based on targeted PCR can allow the simultaneous detection of several taxa, this method being called “DNA metabarcoding”.

DNA metabarcoding was first used in 2011 (Pompanon et al., 2011) to indicate the taxonomical identification of different species thanks to the amplification of metabarcode sequences from eDNA samples. A metabarcode is a taxonomically significant DNA region flanked by two conserved sites functioning as primer anchors for the PCR amplification (Taberlet et al., 2018). After the PCR step, the amplicons are sequenced with a Next Generation Sequences approach (NGS). Different bioinformatic pipelines allow then to match the reads to a reference database in order to generate the list of species included in the eDNA sample analyzed.

In the last years, this tool has been brought to the attention by many studies in ecology and conservation biology (Beng & Corlett, 2020; Vermeulen and Koziell, 2002). The rise in popularity of this molecular technique can be referred to the fact that it is a non-invasive (Maruyama et al., 2018; Schwentner et al., 2021), sensitive (Tucker er al., 2016; Lopes et al., 2017; McColl-Gausden et al., 2021) and cost-effective method (Larson et al., 2020) for faunistic, fungal, microbial and floral monitoring in a large variety of environments.

The analysis of eDNA has shown lots of potentiality especially when the target is a single species (Xia et al., 2021), for example to detect rare species or invasive species, mostly in aquatic

ecosystems (Piaggio et al., 2014; Jerde et al., 2011; Pfleger et al., 2016; Tréguier et al., 2014, Riascos et al., 2016)

This technique is used in a large variety of environment and ecosystems. In terrestrial habitats, metabarcoding has been used to detect both metazoans (Allen et al., 2021), plants (Hartvig et al., 2021) and also to characterize fungal community during ecological restoration (Yan et al., 2018).

Metabarcoding has also been applied to arthropod identification of samples collected using Malaise traps, helping the decision-making process in conservation planning (Ji et al., 2013).

Cave sediments, permafrost and soil samples are employed in paleobiological reconstructions using eDNA; in fact, ancient DNA helped to reconstruct Holocene and Pleistocene plant and animal communities (Willerslev et al., 2003; Epp et al., 2012).

So far, many studies have started to outline standardized protocols especially in aquatic ecosystems (Shu et al., 2020; Loeza-Quitana et al., 2020; Minamoto et al., 2021, Wang et al., 2021). In aquatic habitats, thanks also to the availability of well tested universal metabarcodes (Zimmermann et al., 2011; Leray et al., 2013; Miya et al., 2015; Valentini et al., 2015), many works verified the reliability of eDNA metabacoding in the survey and assessment of fishes (McDevitt et al., 2019), amphibians (Pilliod et al., 2013, Valentini et al., 2016), diatoms (Apothéloz-Perret-Gentil, et al., 2017) and invertebrate communities (Nguyen et al., 2020).

Besides biodiversity investigations, the analysis of aquatic eDNA allows also to assess the environmental impact of offshore gas platforms (Cordier et al., 2019), the ecological quality status of the water in particular using foraminiferal eDNA metabarcoding (Cavaliere et al., 2021) and gamete detection of aquatic species in reproductive biology studies (Bayer et al., 2019).

2.2 The ecology of eDNA

Understanding the origin, state, transport and persistence of eDNA in the environment is crucial for every experimental design. Here, all these aspects will be described and analyzed in a general

eDNA context. Their specific role and influences in freshwater ecosystems will be taken into account in paragraphs 2.3 and 2.4 from the sampling strategy to the analysis and interpretation of the results.

2.2.1 Origin

Despite the increasing interest in eDNA studies, the exact biological and physiological origins of the genetic material gathered as eDNA are still unsure (Barnes & Turner, 2016).

Thanks to the fist researchers, mostly microbial biologists (Ogram et al., 1987), we can state that eDNA can be present in both intracellular and extracellular forms. All living beings possess genetic material, so when it is present in any medium it may come from any single-celled organisms such as Archaea, Bacteria or eukaryotes like Protozoa, Algae or Fungi. Not only the genetic material can derive from active cells but also from dormant or resistant stages such as spores.

Extracellular eDNA can also come from the breakage and release of DNA from cells and tissues that multicellular organisms had shed or dispersed in the environment. Also in this case, the material can derive directly from meiofauna (Tardigrada, Nematoda, Rotifera, etc.), from fragmented organisms' parts such as roots of plants or fish scales. As above reported, the genetic material can derive from active or inactive forms like seeds, pollen, pupas, cysts, etc. (Taberlet et al., 2018).

Several studies pointed out that eDNA present in aquatic and terrestrial environment could develop from fecal materials (Martellini et al., 2005; Andersen et al., 2012). Other body excretions and fluids are source of eDNA such as the mucus of fishes and amphibians (Ficetola et al., 2008). Gametes (Bayer et al., 2019), dead bodies, carcasses and feces (Merkes et al., 2014; Ratsch et al., 2020) are also recognized as eDNA source.

2.2.2 State

It is likely that, with time, the genetic material released in the environment in the intracellular state passes in the extracellular form due to, for example, membranes degradation.

In the different environments, extracellular eDNA can be free or dissolved, particulate in soil, solubilized in water, or linked to the surfaces of organic or inorganic particles (Levy-Booth et al., 2007; Pietramellara et al., 2009).

The work of Turner et al. (2014) highlighted the connection between the possible state of eDNA and the correct method of target eDNA capture. A more comprehensive knowledge of the size distribution of target and non-target eDNA could be crucial in the choice of the best filter size and sampling strategy when dealing with eDNA from water samples (Barnes & Turner, 2016).

2.2.3 Persistence

The destiny of eDNA is influenced by three classes of factors: DNA characteristics (length, conformation, membrane association, etc.), abiotic factors (pH, temperature, light, oxygen, salinity, etc.) and biotic factors (extracellular enzymes, microbial community, etc.) (Barnes & Turner, 2016).

For example, high temperature and microbic activity can shorten the persistence of DNA, whereas if the DNA is associated with particles these could increase its presence in the environment (Taberlet et al., 2018). One of the main elements that could determine the amount of eDNA in a given habitat is the presence and quantity of endogenous nucleases (Hebsgaard et al., 2005). Low temperature may slow the DNA degradation process also lowering the activity of this enzyme (Zhu, 2006).

One of the main concerns in aquatic environmental samples is the DNA hydrolysis, this process being the main cause of DNA damage and breakage in water (Lindahl, 1993). After the collection, all samples containing water should be dried or filled with alcohol to arrest this process.

Another key factor affecting the persistence of eDNA is oxidation. Ionizing radiation may cause DNA modification by direct interaction or may induce the creation of free radicals in the water which can interact with DNA and so damage it (Höss et al., 1996). Pilliod et al. (2014) demonstrated that eDNA was no longer detected in samples exposed to direct sun light after 8

days, whereas it was detected in 20 % of shaded samples after 11 days and 100% of refrigerated control samples after 18 days.

Therefore, the knowledge of the habitat (especially the ecology of the investigated environment) can give a lot of information about the fate of DNA.

In aquatic environments different studies have demonstrated that eDNA presence may vary from few hours to a month (Pilliod et al., 2014; Dejean et al., 2011; Deere et al., 1996). Usually, the presence of DNA can differ also within the water column, the degradation process being higher in the epilimnion and lower in the hypolimnion (Matsui et al., 2001). In running waters, eDNA presence is shorter than in marine environments due to removal by constant currents; it can be therefore used as a present picture of the biodiversity in that habitat (Pilliod et al., 2014).

DNA can last over years in temperate soils as showed by Yoccoz et al. (2012). In colder habitats eDNA can persist up to half a million years in permafrost sediments, so that in particular conditions ancient DNA can give insight about paleobiological biodiversity (Willerslev et al., 2003, Bellemain et al., 2013).

Understanding how long eDNA can persist in the environment is a key factor in order to delimit the time window of the detected organisms.

2.2.4 Transport

Once released, the extracellular and intracellular DNA move in the environment, thus this could affect the detection of species in eDNA investigations. The transport of eDNA is especially an issue in water systems, such as marine and freshwater environment. The genetic material can be shifted from adjacent sites both through active (e.g., biological or anthropogenic transport) or passive (e.g., resuspension, diffusion, leaching) dispersal (Taberlet et al., 2018). Consequently, in streams a species can be detected far downstream from the actual release source of genetic material, depending on the hydrogeological features of the water course (Foppen et al., 2011; Deiner & Altermatt, 2014; Pilliod et al., 2014). In lentic systems this problem is a minor issue even

if some eDNA can arrive in the water bodies from the outside through leaching or settling (Turner et al., 2015; Parducci et al., 2015; Mychek-Londer et al., 2020).

In terrestrial ecosystems, the transport of eDNA seems to be of minor concern; eDNA extracted from soil samples are likely to be linked to organisms living below the ground whereas the contribution of organisms on the surface is minor (Taberlet et al., 2018).

2.3 Freshwater ecosystems

Freshwater habitats are one of the most investigated environments in eDNA studies. The reasons for this are the most disparate. First of all, DNA is a molecule soluble in water: it can therefore circulate and spread in a wide area from the starting origin increasing the likelihood of extracting eDNA from this kind of medium. Another encouraging factor is that the biodiversity assessment in aquatic habitats is becoming mandatory by law in many States (e.g., European Council 2000; Taberlet et al., 2018). Finally, by now, the efficiency and reliability of this method in comparison to the traditional ones are becoming more and more accredited (Civade et al., 2016; Valentini et al., 2016; Pilliod et al., 2013; Shaw et al., 2016; Deiner et al., 2017; Eiler et al., 2018).

Numerous factors could affect the detection of eDNA in aquatic habitat such as biological origin, persistence, diffusion and detectability, as above reported (Taberlet et al., 2018). The taxa which contribute more to the dispersal of genetic material in water are fishes (Civade et al., 2016; Thomsen et al., 2012; McDevitt et al., 2019), amphibians (Valentini et al., 2016; Hobbs et al., 2019; Eiler et al., 2018; Rees et al., 2014) and mollusks (Mychek-Londer et al., 2020; Stoeckle et al., 2021; Clusa et al., 2016). DNA released from arthropods is likely more difficult to detect since they have a resilient exoskeleton made of chitin (Taberlet et al., 2018). The superficial layers of the cuticle are not therefore subjected to cell exfoliation. Both from controlled experimental studies and in field research, it is demonstrated that in water DNA could persist from a few days to a couple of weeks (Dejean et al., 2011; Pilliod et al., 2014), moreover, the permanence time seems to be higher with lower temperature, absence of sunlight and particles decaying DNA (He et al., 2015). Given the relatively rapid degradation of DNA in water, a positive detection is highly

linked with the simultaneous presence of the organism. Especially in lotic systems, DNA can be transported far away from the origin sources and therefore a positive detection not always implies the actual presence of the target species in the sampling site. In running water long distance diffusion is a well-known issue and should be always considered (Jane et al., 2015).

2.3.1 Ditches and canals

Ditches and canals are man-made channels for water. The main difference between the two is the dimension, canals being usually wider and larger than ditches. They are built in order to regulate water levels according to irrigation regimes, or also to connect rivers and stream. Ditches are connected to dewatering pumps or sluice gates to control water levels. During dry periods, the water can be pumped into the ditches from external canals. Since they are often interconnected with each other, aquatic organisms could migrate and move within canals systems. Knowing the water regime in ditches and canal plays a key role in the detection of eDNA because lower water level implies less dilution whereas higher water level could exacerbate the detectability of certain species.

Even if canals and ditches could represent transition zone between habitats creating interesting biodiversity spots in Italy and even if many of them are included in the Directive Habitat 2000, no studies using the eDNA metabarcoding technique are available so far for this geographical area.

Even at a global scale there are few eDNA works on this type of habitats. Thomsen et al. (2012) validated the metabarcoding of eDNA in streams, whereas Mcdevitt et al. (2019) used as a case study an English canal comparing eDNA metabarcoding with traditional surveys based on electrofishing techniques encouraging the use of eDNA as an effective tool to monitor fish biodiversity. Kamimura et al. (2018) investigated fish diversity in a Japanese canal and port area using eDNA detecting more species compared to traditional methods. Jarde et al., (2011) showed the efficacy of environmental DNA as a detection tool for two species of Asian carps in canals and waterways of the Chicago area.

2.4 Sampling

For water samples, researchers tend to use two different sampling approaches based on the presumed concentration of the genetic material in the water. The first one, also chronologically speaking, is the precipitation using alcohol or centrifugation (Ficetola et al., 2008). Ethanol is a DNA preservative, so samples can be stored until the extraction step. Since this method needs low sample volumes, detection of target DNA could be an issue. For this reason, in most of the studies the best strategy has been shown to be the filtration of different volumes of water. The detection probability using this method could be higher since larger volumes of waters tend to be filtrated. However, no agreement has been yet reached upon this thread. In fact, the filtrated volumes could vary from 50 ml (Thomsen et al., 2012) to 100 l (Valentini et al., 2016). Muha et al. (2019) demonstrated that, when working with small water volumes, the filtration using 100 ml syringe is more practical and could gain higher amplification efficiency compared to filtration of larger volumes.

Filtration can be performed either in the field or in the laboratory. A pre-filtration step can be also performed in field and followed by another filtering step in laboratory. An early filtration of the sample may prevent DNA degradation. On the other hand, filtration in a sterile environment could avoid field contamination. Especially if the water sampled is not clean, laboratory filtering is preferred using sterile bottles to carry the samples in dark, refrigerated conditions (Shu et al., 2020). In this case, filtration should be performed within 24 hours after the collection of the field samples (Jerde et al., 2011).

A wide variety of filtering devices, type of membrane and pore size have been used in freshwater eDNA metabarcoding studies. Several types of filters are used to collect DNA from water (Shu et al., 2020) either made of glass fiber (Jerde et al., 2011; Simmons et al., 2016), cellulose acetate (Takahara et al., 2015), cellulose nitrate (Levi et al., 2019; Hänfling et al., 2016), mixed cellulose acetate and nitrated (Li et al., 2018), polycarbonate (Evans et al., 2016; Olds et al., 2016), polycarbonate track-etched (Strickland & Roberts, 2019), polyethersulfone (Xu et al., 2018), nylon (Turner et al., 2014). Spens et al. (2016) proposed an optimized protocol for the extraction

of aqueous eDNA using enclosed filter capsule which can also significantly reduce contaminations. Even if glass fiber filters are the most common one (Shu et al., 2020), it is always recommended to conduct some test to establish the most suitable device.

2.4.1 Storage and preservation

Samples containing genetic material removed from their original context continue to decay after been sampled; for this reason, it is always suggested to keep samples refrigerated especially if the isolation step could not be carried in field right after the collection. In order to stop or at least reduce all the physical and biochemical activities inducing DNA degradation, samples should be stored at least at -4°C following collection. This step could be also done in the field, transferring all the samples in cool boxes containing ice; once arrived at the laboratory, they can be kept at -20°C or -80°C depending on laboratory equipment availability and length of term storage (Lear et al., 2018). Moreover, preserved enclosed filter capsules filled with buffer like ethanol can be stored at room temperature for at least 2 weeks without significant DNA decay (Spens et al., 2016).

Another option is to use buffers that stabilize DNA such as sodium acetate and ethanol (Ficetola et al., 2008) or Longmire's lysis buffer (Longmire et al., 1997; Wegleitner et al., 2015).

Samples like soil, sediment or filter can be completely dehydrated using silica gel in hermetical containers (Shehzad et al., 2012).

2.5 DNA isolation

From the literature analysis regarding eDNA metabarcoding studies, a wide plethora of DNA isolation protocols are nowadays available, mainly depending on the different sample medium and concentration method.

Since it is highly recommended to use the same PCR parameters during the amplification step, one key point is to standardize the DNA extraction for all the samples (Taberlet et al., 2018).

Internal extraction controls could always help as also performing technical replicates could help overcome any heterogeneity issues.

Since soil and fecal environmental samples tend to be extracted with PCR inhibitors, different commercial kits and protocols taking into account this issue are already available (Philippot et al., 2012).

Small volumes of water freshly sampled and not filtrated can be centrifugated to precipitate DNA and perform a classical extraction using commercial kits, salt-extraction using NaCl, conventional cetyltrimethylammonium bromide (CTAB) or phenol-chloroform-isoamyl alcohol. However, it is more common to extract DNA once the water samples have been filtrated (Zhu et al., 2020). DNeasy Blood and Tissue Kit (Qiagen, Hilden, Germany) is a non-hazardous, simple and cost-effective kit often used to extract eDNA from freshwater samples. A DNeasy Blood and Tissue modified protocol by Spens et al. (2016) combined with Sterivex-GP polyethersulfone, pore size 0,22 µm enclosed filter yielded higher amounts of total DNA compared to other 41 different treatments.

2.6 PCR and multiplexing

The PCR is a technique used in molecular biology to amplify a target segment of DNA generating thousands of millions of copies. This method relies on thermal cycling where the reagents are exposed to cycles of repeated heating and cooling, allowing different temperature-dependent reactions, such as DNA melting, primer annealing and DNA replication, to proceed many times in cycle. The target DNA sequence is selected by the use of specific primers, short synthesized oligonucleotides, containing sequences complementary to the target flanking regions; these, along with a DNA polymerase, enable selective and repeated amplification. For each cycle, the newly synthesized DNA fragments are also used as a template for further replications. In this way, the original target DNA is exponentially and continuously amplified.

Within each PCR cycle, the first step is called denaturation and it consists in the DNA melting; the high temperatures, about 94-95 °C, induce the double-stranded DNA to separate by breaking the hydrogen bonds between the two strands.

In the next step the temperature is lowered, thus allowing the annealing of the primers. Forward and reverse primers are generally included in the reaction mixture, one for each of the two single-stranded DNA target region. The efficiency and specificity of the whole reaction is critically defined by the temperature of annealing, in fact it has to permit the hybridization of the primer to the strand.

The third step is the elongation or extension phase. In this step, the DNA polymerase, typically a thermostable polymerase, synthesizes the complementary strand by adding the dNTPs from the mixture in the 5'-to-3' direction. This procedure is repeated 30-40 times (cycles) in each PCR reaction.

One of the most limiting factors in PCRs involving environmental DNA is the length of the amplicons. In fact, in routine PCRs it is easy to amplify even up to 1000 bp, while, if the template DNA is highly degraded as usually in environmental or ancient samples, the amplicon size should not go beyond 150-200 bp since the starting material is represented mostly by short fragments. Dealing with larger amplicons implies the adjustment of PCR settings, library preparation and the choice of the right Next Generation Sequencing (NGS) platform.

Another issue involving DNA metabarcoding is the production of false positive. This can happen, for example, when contaminant genetic material is amplified during the PCR step. This not only influences the amplification step but also the sequencing step, leading to an overestimation of the number of taxa (Coissac et al., 2012). To overcome this issue, new enzymes are being synthesized very often such as proof-reading polymerases. Even if, theoretically, the same Taq polymerase and PCR conditions should be used in all DNA metabarcoding studies, the different sample media, primers and groups of interest do not allow it (Lear et al., 2018). That being said, it seems that hot start Taq polymerases could be a good compromise if we consider quantitative

aspects, specificity and costs (Taberlet et al., 2018). During a hot start PCR, an antibody blocks the Taq activity at low temperature so an initial step at 95-96°C is needed to activate it. In this way nonspecific priming and primer dimer formation during the first increase of temperature are avoided.

In order to tackle the stochasticity of PCR, especially given all the precautions and arrangements with environmental DNA, it is recommended to perform different PCR replicates. The number of replicates should be adjusted to the study design. More replicates mean that more elusive or not so well represented taxa can be detected.

2.6.1 PCR controls

To correctly analyze correctly metabarcoding results and detect and monitor possible false positives and potential contaminations it is crucial incorporate appropriate controls in the PCR workflow.

Extraction negative controls or blanks consist in a extraction done during the regular extraction with the exact same reagents, but the DNA source. The inclusion of this kind of control is strongly recommended to check for contaminations. The identification of contaminants during the extraction could be achieved comparing the results from the extraction negative controls and the PCR negative controls (see below). Contamination is a concerning issue when dealing with diluted samples or universal primers since contaminants may derive not only from field or laboratory processes but also commercial kits can contain traces of them (Leonard et al., 2007).

Negative PCR controls correspond to a PCR mix where the DNA template is replaced with DNA-free water used for the dilution of the reagents. In this way the absence of contaminants in the amplification reagents could be checked, even though the absence of amplification in such controls does not often entirely correspond with the actual absence of contaminant due to the so-called carrier effect (Kolman & Tuross, 2000). Low concentrated contaminant DNA can stick to plastic tubes and tips and for this reason it will not be amplified; but once DNA from samples is added this could compete for adherence and make the contaminant DNA available for the

amplification (Malmström et al., 2005; Leonard et al., 2007). Filtering strategy during the bioinformatic step can be set to discard these sequences.

Positive PCR controls are mock communities. They should comprise taxa related to target group but from other region/ecosystems and they should have the same DNA concentration. These types of controls help to ensure the PCR ability to amplify target sequences.

Finally, if a single-step PCR with tagged system is employed for multiplexing samples (see below), it is recommended to comprise some unused tag combinations. During bioinformatic analysis it is possible to identify chimeras and filter sequences assigned to the unused tag pairs (Schnell et al., 2015)

2.6.2 Multiplexing samples

The utilization of any NGS platforms to analyze sequences from environmental samples generates an extremely high quantity of raw data as a consequence of the high sequencing outputs. In order to benefit from all the advantages, time and costs of this technology it is recommended to pool all the samples into a single run to retrieve all the raw sequences from all the samples in a single time. This process is referred to as multiplexing of samples in a single DNA sequencing run.

It should always be taken in consideration that all the NGS platforms have been firstly designed for genomic DNA analysis, hence most of protocols are not thought and enhanced for PCR amplicon sequencing. Here the structure of an amplicon (Fig. I1) that can be sequenced on Illumina platforms will be described as an example to better understand how multiple samples can be sequenced in a single run.

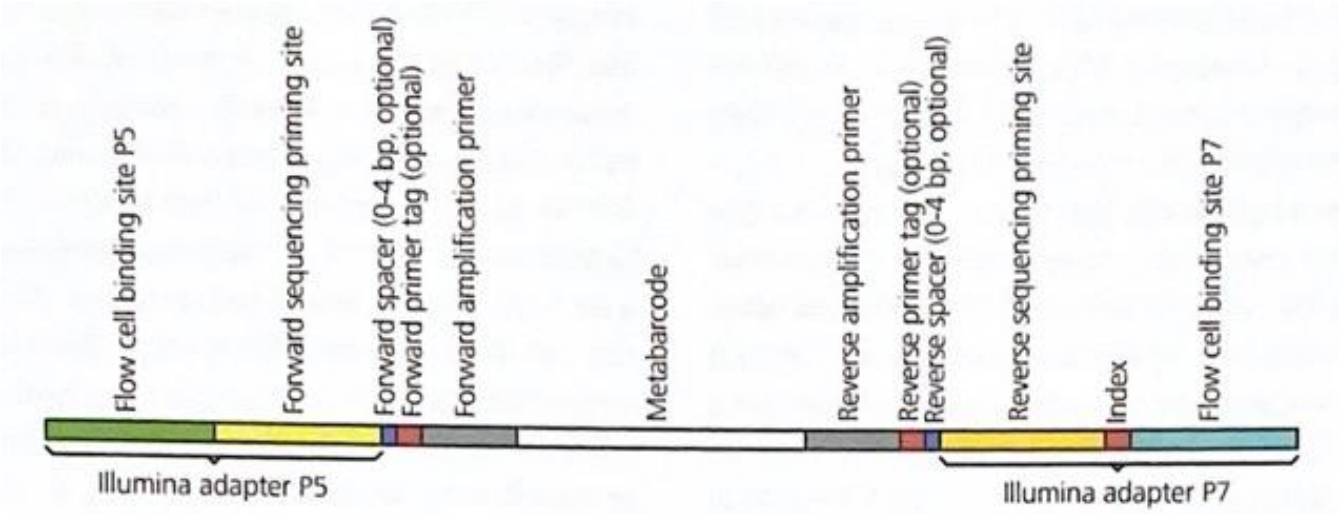


Figure I1. Structure of an amplicon that can be sequenced on Illumina platforms (Taberlet et al., 2018)

The metabarcode in this case is flanked by the Illumina adapter P5 on one side and by the Illumina adapter P7 on the other. The external part of the adapters allows the flow cell binding since they are complementary to oligonucleotides attached to the flow cell. The internal part consists in the forward and reverse sequencing priming sites. The two regions of the adapter P7 are separated by a 6/8 base index.

The first possible strategy (Fig. I2) relies exactly on using the Illumina indices and consist in a single PCR with Illumina adapters. In the amplification step, long primers (usually 75-95 bp) are used and they are formed by the taxa primer and the Illumina adapters containing the indices on the 5'-end. This strategy is simple: the amplicon is ready to be loaded on the NGS platform and if no tags are used there will not be tag jumps (Schnell et al., 2015) since this issue occurs during the library preparation. The number of sample tags could be implemented if an index on the P5 adapter is added (double indexing). The first concern is that this strategy is effective only if the genomic material is highly concentrated due to the lower mobility of long primers. Even if tag jump is not an issue, index jump is still a problem during the sequencing steps. A read could be assigned to a different index if, during the sequencing of the index, the signal of a cluster is stronger than the signal of a weaker nearby cluster.

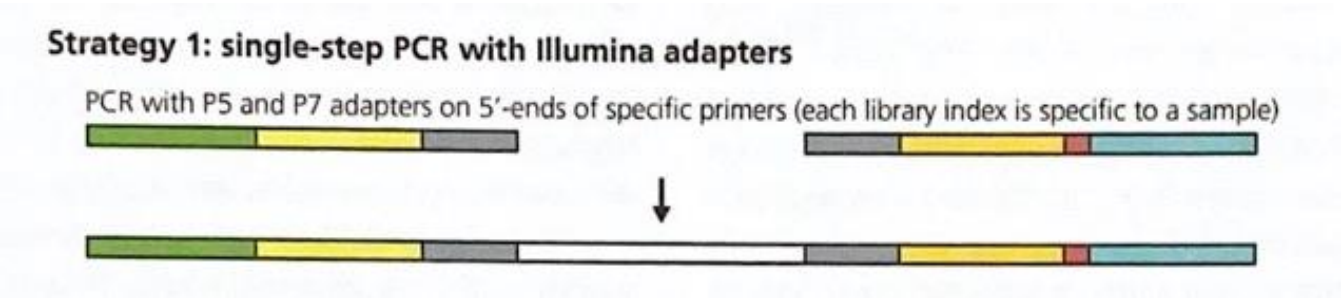


Figure I2. First strategy for multiplexing several samples on the same sequencing lane of Illumina sequencing platforms (Taberlet et al., 2018)

The second strategy (Fig. I3) consists in a two-step PCR with Illumina adapters. The first PCR is carried out with the specific taxa primers linked at the 5'-end with the forward or reverse sequencing priming sequence of the Illumina adapter; the second one adds the flow cell binding sites and the indices. Also in this case the double indexing is possible (Fonesca & Lallias, 2016). Compared to the first strategy, the success of amplification with low DNA concentration is higher since the length of the primers is reduced to 40-45 bp. Another advantage is that the second couple of primers could be reused for marking other amplicons in other barcoding experiments, if the sequencing priming regions tailed to the primers remain the same. The first disadvantage is the higher contamination rates involved in two-step PCRs. As the first strategy, even with double indexing, the number of different tagged samples is limited. Lastly, also this approach is subjected to the index jump problem.

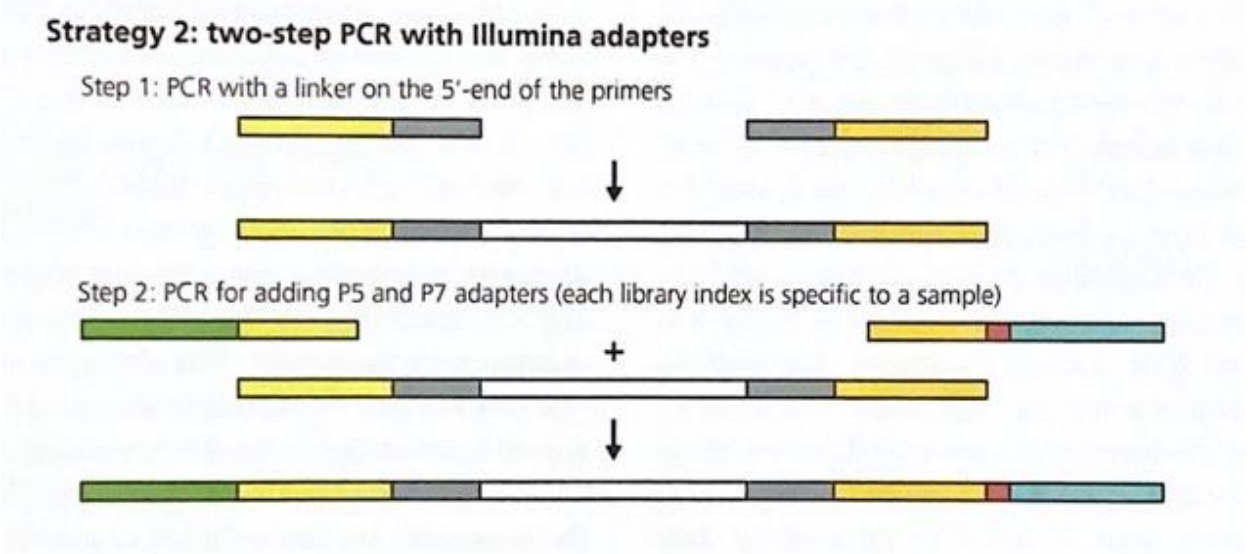


Figure I3. Second strategy for multiplexing several samples on the same sequencing lane of Illumina sequencing platforms (Taberlet et al., 2018)

The third possible (Fig. I4) strategy is a single-step PCR with tagged primers. A single PCR is performed using short tagged primers, allowing also the detection with low concentrated DNA templates unlike the two previous approaches. There are two different ways to tag the samples. One could use all different tag combination so that to a sample corresponds one single tag combination. Using tag combination reduces the cost and allows the tag jumps detections. This could be achieved if during the experimental preparations some tag combinations are not used (tag blanks, see above). The second possibility is to use all different tags to mark the samples, in this way tag jump is unlikely to happen but it is a less cost-effective method. With this strategy it is also possible to increase the quality of the base calling if 2 to 4 random nucleotides are added at the 5'-ends of the tagged primers. If this strategy is chosen, the library preparation should not involve any PCR cycles or T4 polymerase step in order to lower the occurrence of tag jumps. The MetaFast protocol by Fasteris SA is specifically designed to limit this issue. Finally, also with this strategy the index jump is still a problem, but it can be managed loading a different index for each metabarcode in the same sequencing lane or if different libraries with the same metabarcode have to be sequenced, they can be loaded in different lanes.

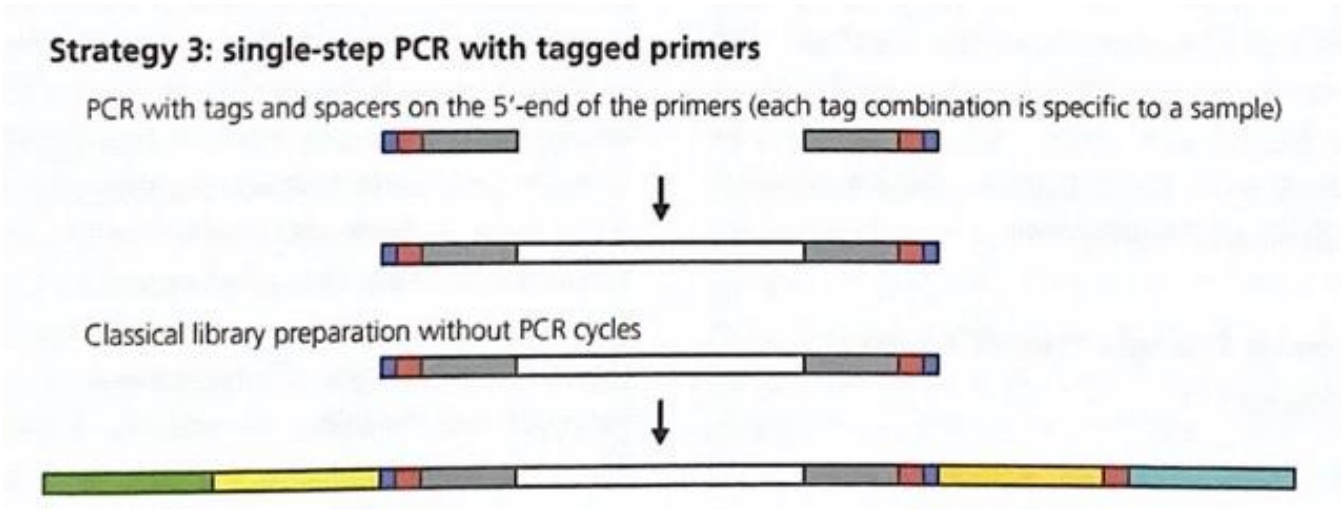


Figure I4. Third strategy for multiplexing several samples on the same sequencing lane of Illumina sequencing platforms (Taberlet et al., 2018)

2.7 Sequencing

The majority of amplicon sequencing from eDNA metabarcoding studies is made using Illumina platforms. The costs could vary a lot among the different application since there are available lots of analysis options such as different sequencing chemistry and multiplexing options.

Illumina, since the first time it was introduced in the sequencing industry in 2007 (Shokralla et al., 2012), has continued to manufacture different platforms with multiple sequencing outputs and read length. Surely, it became popular thanks to more affordable cost per sequence (Buermans & den Dunnen, 2014), high throughput and large coverage. Another great advantage is the possibility to perform paired-end reads. Paired-end sequencing allows to sequence both ends of a fragment and to generate high-quality, alignable sequence data. In addition to producing twice the number of reads at the same time and effort in library preparation, sequences aligned as read pairs enable more accurate alignment and the ability to detect indels.

One of the most known concerns in sequencing amplicons with Illumina sequencers is linked to the fact that, since they are built to sequence genomes, the machine expects the same proportion of A-T and C-G and AT/CG ratio during each sequencing cycles and this is almost never the case with PCR amplicons. The base calling is adjusted on the first 25 sequencing cycles, but since in amplicons these ratios are different the base calling algorithm is compromised. In order to deal with this problems a couple of solutions are available. First of all, one strategy to improve the base calling involves the improvement of the sequence diversity by adding genomic DNA from the phage PhiX with the amplicons (Kozich et al., 2014). An additional strategy relies on phasing the amplicon sequencing (Wu et al., 2015) by adding a number of additional nucleotide bases at the 5'-ends of the primers (as described in the third multiplexing strategy).

2.8 Bioinformatic pipeline

Classical ecology and biodiversity census have to face with organized sampling campaigns aimed at retrieving and reconstructing a species inventory as much as complete and representative it can be. One of the main challenges shared by all the analyses performed using any NGS platforms

is the management of the of huge amount of data produced. Classical sequences analysis, such as the one performed to handle data from Sanger technology, cannot be employed in eDNA metabarcoding data analysis. This kind of data elaboration needs a large amount of CPU and RAM power, so it should be performed on dedicated servers using UNIX systems and not on a simple personal computer. The most critical issue dealing with metabarcoding data is its "noise". In fact, along with the biologically significant data also lots of artifacts are present in a metabarcoding analysis such as PCR and sequencing errors, chimeras, and so on (Zinger et al., 2019; Schnell et al., 2015). These biases can be attenuated thanks to the implementation of appropriate filters during the bioinformatics analysis using different sets of programs. Even if several packages and programs are available, no agreement on which pipeline or set of programs to use among researchers in the eDNA metabarcoding field has been reached (Bazinet & Cummings, 2012; Prodan et al., 2020; Sczyrba et al., 2017; Siegwald et al., 2017). The final bioinformatic step allows to clean the data and finally produce a species list. Even if bioinformatic pipelines can differ very much one to another, six fundamental steps are pretty much always present: paired-end read pairing, demultiplexing, dereplication, quality filtering, removal of errors, taxonomic assignment (Mathon et al., 2021). These steps are common in every metabarcoding data management, but they might differ in order and tool used depending also on the differing sequencing and multiplexing approaches. Nowadays, several pipelines are available to analyze in a rather straightforward way the raw data and obtain the final species lists. The most common and used are OBITools (Boyer et al., 2016), QIIME2 (Bolyen et al., 2019), VSEARCH (Rognes at al., 2016). Some tools could be excluded *a priori* because they are not compatible with the data produced; for example MOTHUR (Schloss et al., 2009) is only suited for the *16s* rRNA sequence analysis while Kaiju (Menzel & Krogh, 2016) for protein-level assignment.

Even before the start of the pipeline, once received the raw data a global quality determination of the sequencing run should be checked. If the global quality of the raw data is too low, a resequencing of the samples should be immediately taken into consideration. FastQC (Andrews, 2010) is the standard quality control tool for high throughput sequence data. The output from

FastQC is an html file which contains one result section for each FastQC module. Aside from the graphical or data list provided, a flag of “Passed”, “Warn” or “Fail” is assigned. One should be cautious with these flags since they are assigned on very specific assumptions applicable only to good quality whole genome shotgun DNA sequencing: they are less reliable with targeted amplicon sequencing. Expected results for this kind of analysis are: biased per base sequence content, narrow distribution of GC content, very high sequence duplication levels and extreme abundance of overrepresented sequences.

2.8.1 Paired-end read pairing

The first bioinformatic step consists in the merging of the forward and reverse reads on their 3’-end to form a single consensus sequence. The command *illuminapairedend* in the OBITools uses an exact alignment algorithm which considers gaps and also the possibility of having a metabarcode shorter than the reads length or longer than a paired-end overlap. It considers the quality score of the reads at all positions during the alignment so that the quality of the consensus is maximized. PEAR (Zhang et al., 2014) operates in an equivalent way, on the other hand, in FLASH (Magoč & Salzberg, 2011) the quality scores are not considered.

2.8.2 Demultiplexing

Since many PCR products from different samples are pooled together in a single library, each full amplicon sequences can be assigned to its sample using the specific tag combination employed in the multiplexing step. In the same moment, also the primers are removed. The two main programs that can perform this step are cutadapt (Martin, 2011) and *ngsfilter* from the OBITools package.

2.8.3 Dereplication

Amplicons from metabarcoding experiment sequenced using NGS platforms occur multiple times. To reduce the size of data, a good practice is to keep only one representative sequence and the count information. This operation is known as dereplication, and it could be seen as a

clustering process with a 100 % identity threshold. This step can be executed in VSEARCH or in OBITools with the command *obigrep*.

2.8.4 Quality filtering

Sequences can be inspected so that artifact and errors can be removed. All metabarcodes containing ambiguous bases can be removed, in this way only sequences with A, C, G, T are kept. Amplicons with a length differing from the expected size can be discarded as well. The *obigrep* command from OBITools allows filtering based on a large set of option and sequences attributes. VSEARCH and cutadapt too permit to filter sequences shorter or longer than a certain value or/and ambiguous bases.

This step is crucial for removing chimeras, that are any sequences formed when two or more sequences are combined together during PCR. In particular, they are generated when incomplete extension occurs in one round of PCR and then the resulting fragment acts as a primer for a different sequence in the next round. Chimeras are a problem because they can artificially alter the diversity estimates, so they need to be removed in these filtering steps. Chimera removal can be performed using chimera-free reference databases, as with Chimera-Slayer (Haas et al., 2011) and DECIPHER (Wright et al., 2012). Whereas UCHIME (Edgar et al., 2011) and Perseus (Quince et al., 2011) can accomplish this task creating a *denovo* reference database from the analyzed sequences themselves.

2.8.5 Error removal

The removal of variants and PCR errors is fundamental to group together biologically significant sequences and also to avoid the overestimation of species. There are two different ways for grouping similar sequences together: denoising or clustering. Reads can be denoised into amplicon sequence variants (ASVs) removing or correcting noisy reads or collapsing highly similar sequences into singular molecular operational taxonomic units (MOTUs).

The denoising methods currently available in QIIME2 include DADA2 and Deblur. Both perform internal chimera checking and abundance filtering. These methods filter out noisy sequences,

correct errors in sequences (DADA2), remove chimeric sequences and singletons, join denoised paired-end reads (only DADA2), and then dereplicate those sequences. The features produced by denoising methods are known as “amplicon single variants” (ASVs), “sequence variant” (SV), “exact single variant” (ESV), etc. The denoising step in OBITools can be executed using the *obiclean* command, but since it is very slow and stringent with long amplicons it might remove natural variability so it should be applied for example only when analyzing highly conserved metabarcodes such as *18s* rRNA.

Clustering method can have a constant or a variable identity percentage cutoff threshold. Constant cutoff clustering software such as UCLUST (Edgar, 2010) or sumaclast (Mercier et al., 2013) rely on the barcoding gap which assumes that the metabarcode intraspecific variability is lower that the interspecific one. Since the barcoding gap depends on the taxonomic group analyzed, it is often difficult to determine the correct criterium to choose the cutoff.

Flexible cutoff clustering programs depend on different algorithm. CROP (Hao, Jiang & Chen, 2011) uses a Bayesian clustering algorithm with variable thresholds, denying the barcoding gap explained before. Wide taxonomic ranges can be clustered in just one step and sequences with random errors may be recovered. The main disadvantage is that Bayesian clustering is heuristic, so results are not repeatable. This software has also huge computation times, even with parallel computing, finally if low pre-clustering values are used, CROP leads to the formation of Super-cluster. Swarm (Mahé et al., 2011) is deterministic and faster than CROP; it relies on a step-by-step aggregation algorithm, it forms a network clustering by a distance value which is an index of distance between reads. With any clustering method chosen, a recounting procedure (mapping) is always needed after the cluster grouping. The initial reads are mapped against the obtained clusters and are summed up for each sample.

2.8.6 Taxonomic assignment

The final step of a metabarcoding bioinformatic pipeline is the taxonomic assignment, that is the classification of the curated and filtered sequences into taxonomic classes. The identification most of the time implies the comparison between unknown eDNA metabarcodes and a reference database with taxonomic information. A reference database can be extracted from online resources such as EMBL (http://www.ebi.ac.uk/embl/index.html), NCBI (https://www.ncbi.nlm.nih.gov) or organism-specific database such as BOLD (www.boldsystems.org; specifically for metazoans *coi*, plants *rbcL* and *matK* and the ITS region of fungi). Reference databases can also be built locally using DNA amplification of identified samples with primers of the same metabarcode region that will be Sanger sequenced independently or using NGS platforms.

The query sequences (ASVs or OTUs) are compared to the reference database sequences with known taxonomic annotation. This comparison allows to find sequence similarity between the query and the reference but merely finding the closest alignment is not enough since other sequences equally similar could have a different taxonomic assignment. So, the taxonomic annotation is based on alignment or k-mer frequencies which determines the closest taxon with a certain degree of confidence or consensus.

The OBITools pipeline provides the *ecotag* command which uses a supervised classification algorithm, based on the search of the lowest common ancestor (LCA), that annotates sequences with several taxonomic attributes. The best match to the query sequence in the reference database is found with a given identity percentage. Other sequences in the reference database that are closer or equally close to the best match to the query sequence are also considered so that the query sequence is assigned to the most inclusive taxon comprising all these reference sequences. Sequences which have low identity with any reference sequence in the database, are usually assigned to high-rank taxa.

QIIME2 has two different approaches for assigning taxonomy, the first involves aligning reads to reference databases directly using a consensus approach of taxonomy assignment using BLAST (https://blast.ncbi.nlm.nih.gov/Blast.cgi) and a local alignment (*classify-consensus-blast*) or VSEARCH global alignment (*classify-consensus-vsearch*). The second one uses a machine learning classifier to assign likely taxonomies to reads and can be used in the classify-sklearn method.

Currently, no unanimous agreement has been reached on which taxonomic classification tool is optimal. The choice must always be on experimental design, computational power and time, taxonomic group studied, reference database availability and metabarcode used.

3 A CASE STUDY: SCOLO DOSOLO

Aim of this study is to test the technique of the metabarcoding of eDNA from freshwater samples in an anthropic canal in Italy, in order to assess and characterize the fish and freshwater invertebrate community. No multi-taxa approach studies on freshwater eDNA in Italy are known so far. A preliminary activity has been carried out during June 2019 to test different sampling strategy, filtration devices and filters, wet lab procedures, library preparation and sequencing. The results from this test were also used as a comparison with the data produced in the 2020 campaigns. Subsequently, four different season sampling were carried on during 2020, with two electrofishing campaigns performed in the summer and autumn field surveys. Many complications and limitation due to Covid-19 pandemic impacted on field work at different levels, but on the whole at least all environmental samples have been collected and processed.

3.1 Dosolo

Scolo Dosolo (Figure M1) was chosen as study area and three points along the canal were identified for samplings. The gutter is an anthropic canal located in Sala Bolognese (Bologna, Italy) and it belongs to the Special Protection Area (SPA) Cassa di espansione Dosolo, a Natura 2000 protected area (Habitat Directive; IT4050030). The flood retention basin of Scolo Dosolo is delimitated by Scolo Dosolo itself and Collettore delle Acque Basse and it expands for 62 hectares. In 1993, the area got a hybrid assignment between wood production and environmental restoration thanks to the realization of Area Riequilibrio Ecologico “Dosolo” comprising a 3 hectares central humid region surrounded by 2.5 hectares of hygrophilous reafforestation.

Figure M1. Photo of Scolo Dosolo near site 2.

The management of the surroundings involves the periodic handling of the water levels in the humid areas to guarantee the maintenance of sufficient water levels and the presence of specific habitats (e.g., muddy shoals and rushes).

The lack of such habitats in the nearby valley is the reason why many organisms are attracted to this area making it the perfect study target for this environmental study.

The presence of a waterway plays a strong role in the colonization by animal and plant species. Four habitats of Community interest are present: two stagnant water habitats (showing an annual mesotrophic vegetation (Natura 2000 Habitat Code: 3130) and a perennial and slightly fluctuating vegetation (3150)), and two natural grasslands, a thermoxerophytic one (6210) and a wet hygrophytic the other (6410). Helophytic margins of faunal interest are also present.

No fish species from Habitat Directive are present in the area. In the humid region a well-structured fish community is lacking probably due to the intermittent water availability in Scolo

Dosolo. Moreover, the water quality level itself is low and in the deepest areas the substrate is anoxic.

In Scolo Dosolo the species richness is considered higher in comparison to the surrounding areas.

The main contribution to species richness is given by invasive species, such as the stone moroko (*Pseudorasbora parva*), the crucian carp (*Carassius carassius*) and the mosquitofish (*Gambusia* sp.), these being also the most abundant ones. Even if communities of carps (*Cyprinus carpio*) are present in less quantity, they are well-structured. The common sunfish (*Lepomis gibbosus*) is the invasive alien species showing the lowest density number.

The only two autochthonous species, namely, the European chub (*Squalius cephalus*) and the Italian bleack *Alburnus arborella*, show a low density since they might be suffering the competition with all the invasive species present in a larger number.

Among invertebrates in the canal also the white-clawed crayfish (*Austropotamobius pallipes*) has been reported even if their populations are threatened by predators and diseases. Juvenile life stages are often eaten by aquatic coleopterans or Odonata nymphs, while adults are predated by fishes like *Lepomis gibbosus* or *Squalius* spp*.* The species is classified as endangered by the IUCN red list. The presence of alien species of crustaceans like *Procambarus clarki* means more competition for resources and they are carriers of diseases like the crayfish plague caused by the fungus *Aphanomyces astaci*.

Alongside with few individuals of the autochthonous crayfish, a great population of the red swap crayfish (*Procambarus clarkii*) is present in all the area. All present information is reported in Quadro Conoscitivo (QC) - Cassa di espansione Dosolo (https://ambiente.regione.emilia-romagna.it/it/parchi-natura2000/rete-natura-2000/strumenti-di-gestione/misure-specifiche-di-conservazione-piani-di-gestione/misure-di-conservazione-per-sito/QC530.pdf/@@download/file/QC530.pdf)

3.2 MATERIALS AND METHODS

3.2.1 Sampling campaigns

Three sampling points at 500 m ca. distance (Fig. M2) were chosen for both the traditional (electrofishing) and eDNA approaches.

During the 2019 preliminary campaign, at each location an on-site prefiltering step was performed on a two liters sample and on one liter of PCR-free waters used as a blank field sample. For this field operation, Nalgene™ Rapid-Flow™ Sterile Single Use Vacuum Filter Units (TheromoFisher) with 1 liter capacity and a PES membrane with 0,45 µm pore size were used. Two bottles of 1 liter water each were also filled without prefiltering. All the water samples were then filtered in laboratory under sterile conditions using enclosed filter capsules.

2020 was devoted to the sampling campaign with a winter (January), spring (May), summer (July) and autumn survey (October; Tab. M1). This time two bottles of water (1 liter each) were collected at the three sites along Scolo Dosolo and no prefiltering step was applied since from preliminary results this operation came out as non-necessary as it might introduce contaminations from field and reduce the taxa that could be detected.

After each water sample collection, the Hydrosynergy co-op. staff carried out the electrofishing (see below) to assess the freshwater fish community with the exception of the winter and spring samplings.

It is to be noted in fact that during the winter survey only the eDNA water samples were collected because the traditional freshwater fish community evaluation using the electrofishing was not possible due to the absence of an appropriate water level. We were able to collect only the eDNA samples also in the spring survey; the activity in fact took place during the phase 2 of the coronavirus outbreak so the traditional electrofishing survey was not allowed because of safety precautions; the summer survey saw both the collection of eDNA freshwater samples and the assessment of the fish community thanks to the electrofishing performed by the Hydrosynergy co-op staff. During this field trip, also some insects and water invertebrates were collected using

an entomological net and a surber sampler, respectively to verify eDNA and morphology protocol also on invertebrates. In the last sampling campaign, we were again able to perform both the eDNA samples collection and electrofishing.

Table M1. Date of every sampling session with corresponding survey

Date	Season	eDNA	Electrofishing	Invertebrate Collection
25/06/2019	Summer (preliminary)	x	x	no
16/01/2020	Winter	x	no	no
12/05/2020	Spring	x	no	no
03/07/2020	Summer	x	x	x
20/10/2020	Autumn	x	x	no

Figure M2. Cassa di espansione Dosolo with the 3 sampling sites. The flow direction is South to North.

3.2.2 Authorization

A quite significant amount of time was devoted to request all the authorizations to the competent Institutions to perform, alongside the water sampling, also the traditional faunistic surveys involving the electrofishing technique. The application regarded the following permissions:

- ISPRA (Istituto Superiore Protezione Ricerca Ambientale),
- STACP Bologna (Servizio Territoriale agricoltura, caccia e pesca),
- Servizio Aree Protette Foreste e Sviluppo della Montagna della Regione Emilia-Romagna
- Ministero della salute via Comitato di Benessere Animale of the University of Bologna.

The permits were obtained as follows:

- ISPRA prot. 5317 applied on 15/01/2019 and obtained on 12/02/2019
- STACP PG/2019/212219 applied on 15/02/2019 and obtained on 01/03/2019
- Servizio Aree Protette Foreste e Sviluppo della Montagna della Regione Emilia-Romagna PG/2019/0169024 applied on 15/01/2019 and obtained on 15/02/2019
- Ministero della salute via Comitato di Benessere Animale 446/2019 applied on 10/12/2018 and obtained on 12/06/2019

3.2.3 Traditional Surveys

To compare the metabarcoding data, traditional surveys were performed involving both the fish community and invertebrates. The freshwater fish fauna was investigated thanks the application of the electrofishing technique, while invertebrate community was surveyed using entomological net and surber sampler. According to the regulation (EU) No 1143/2014 on the prevention and management of the introduction and spread of invasive alien species, the Decreto Legislativo 230/2017 and Programma ittico regionale 2019/2020, all the invasive species collected with the

electrofishing technique (i.e. all the species with the exception of *Alburnus alborella* and *Cyprinus carpa*) were detained by the Hydrosynergy co-op. staff members.

Electrofishing

Electrofishing is a rapid and non-invasive ichtyofauna capture method. It is based on the effect induced by electric fields on fishes leading to the capture of individuals of different species/size classes; it is not selective and it allows a wide view on the quality and quantity of fish population present in a specific water stream. The electrofishing device generates an electric field in water between two submerged electrodes, the anode (positive pole) represented by a rod made by an isolated material, equipped with a "human-present" switch with at one extremity a metallic ring handled directly by the operator and the catode (negative pole) made by copper o other material (the tail). The electric field induces a passive attraction effect on fishes towards the net, the so-called electrotaxis, so that they can be caught by the collector thanks to the landing net. The efficiency of this method is higher in sites where the water does not exceed 2 meters of depth and the water conductibility is higher than 100 µS and lower than 700 µS (Scottish Fisheries Co-ordination Centre – SFCC, 2007).

In shallow waters, the survey operations have been conducted as suggested by the APAT 2007 protocol (Sollazo et al., 2007). The electrofishing method has been conducted investigating the riverine area (1-2 meters of depth) using a portable electric stunner FEG 3000 (Power 3 Kw; max output power 500 V) through transects length 20 times the width of the canal always upstream the sampling points.

Invertebrates sampling

Terrestrial arthropods were collected by a transect-wise sweep net sampling (diameter 40 cm, mesh width 1.5 mm) following the protocol by Rauschen *et al* (2009). Sweep netting was carried out along the canal bank near the sampling sites at approximately 5 m from the canal, to avoid edge effects.

Bentic invertebrates were collected using a Surber sampler with 375 µm net and sampling surface of 500 cm^2 along a 3 m transect upstream. All the samples were collected by the canal bank according to the I.B.E. protocol by Ghetti (1997) and APAT-IRSA (2004).

Morphological Identification

All the fish species were identified during the electrofishing operations by fish specialist of the Hydrosynergy co-op. crew. After the identification and the data annotation, a photograph for each species was taken (Fig. S1 – S7).

For the invertebrate morphological analysis, i.e. to build a morphological data set, the insects sampled with the net were pinned and let dry out, whereas the benthos specimens collected with the surber sampler were conserved in 70 % ethanol.

A total of 36 specimens were photographed on a stereo microscope (Fig. S8 – S38) and identified using morphological dichotomic keys.

3.2.4 Traditional Barcoding

Since the identification of the invertebrate specimens was performed by a non-specialist researcher (me) a standard barcoding protocol, to support morphological analyses, was also carried out.

Total DNA was isolated from the dried individuals using the NucleoSpin DNA Insect (Macherey-Nagel) extraction kit and from ethanol preserved specimens using the Smarter Nucleic Acid Preparation (Stratec) extraction kit following the manufacturers' instructions.

Product yield, specificity and potential contamination were monitored by agarose gel electrophoresis. The folmer region of the cytochrome oxidase subunit 1 which is the current sequence used for the metazoan barcoding was then amplified (Vrijenhoek, 1994)

The PCR cycle consisted in a first step denaturation at 95 °C for 5 min, a second step repeated for 33 cycles at 95 °C for 30 sec, 55 °C for 30 sec and 72 °C for 45 sec, and the final extension at 72 °C for 7 min. The 30 PCR products obtained were purified and cleaned using ExoSAP-IT PCR Product Cleanup Regent (Thermofisher) and sequenced by MACROGEN Europe with the BigDye Terminator chemistry, based on the Sanger method.

Electropherograms were screened through the program SeqTrace; after this step, 21 sequences resulted suitable for the downstream analysis.

The final step consisted in the taxonomic assignation from BOLD and the results were compared to the traditional assignment using the morphological identification keys.

3.2.5 eDNA Laboratory Methods

Samples

The water samples were collected in 2 sterilized 1000 mL Nalgene bottles (Thermo Fisher Scientific Inc. USA) for each site. Prior to and after the sampling they were wiped and rinsed with 5 % bleach and then with laboratory-grade ethanol (70%). Before each water collection, the bottles were soaked in canal water away from the sampling point so as to remove ethanol and bleach residuals. Nitrile gloves were used during cleaning, collection, and all laboratory procedures. All the eDNA samples were collected before the traditional surveys and always going from site 1 to site 3, that is from downstream to upstream in order to avoid contamination from site to site.

After the collection, all the samples were kept refrigerated in a cooler bag and then carried to the MoZooLab (Department of Biological, Geological and Environmental Sciences, University of Bologna) and kept in fridge until the filtration step which occurred within 24 hours from the collection.

Lab conditions

All the pre-PCR, PCR and post-PCR steps took place in different and separated rooms. Prior to any lab procedures, all surfaces were cleaned with bleach and molecular-grade ethanol (70%). Filtration, eDNA extraction and PCRs were carried out in laminar flow hoods which were UV-treated before and after, alongside all the laboratory instruments. Gloves were worn at all lab steps and changed between the handling of each sample.

Filtration

All the samples, also the prefiltrated on site during the preliminary study, were filtrated (40-70 ml) until the filters were blocked using 'enclosed' filter capsule (Sterivex-GP polyethersulfone, pore size 0.22 um (Merck; ww.merckmillipore.com). UV-treated molecular grade ethanol 99% was used as storage buffer, all filters were stored at room temperature until the extraction step which occurred within one week. One blank consisting in PCR-free water for each season sampling was also analyzed.

Extraction

DNA extraction was carried out from the emptied filter within the capsule and from the preservation buffer within a centrifuge tube following the protocol from Spens et al. 2017.

For all the downstream applications, Eppendorf DNA LoBind tubes were used in order to avoid DNA retention by plastic, a well-known issue documented for short DNA fragments (Ellison et al. 2006; Gaillard & Strauss 1998).

The modified protocol proposed by Spens et. Al. (2017) using the DNeasy Blood & Tissue kit (Qiagen) was followed to carry on the extraction from both the buffer and the filter, which were treated as technical replicates.

For the preliminary study, a total of 18 sample were extracted: 1 prefiltered sample, 1 not prefiltered, 1 field blank for every sampling site both from the buffer and the capsule.

In total, 32 samples were extracted from tree sites in Scolo Dosolo (1 sample of water per site per season, 1 blank for each campaign both from buffer and capsule) during the 2020 campaign.

DNA quantification

DNA total yield was quantified using a Qubit 1.0 fluorometer (Thermo Fisher Scientific Inc.) applying the BR assay for dsDNA (Life Technologies).

Primer choice

During the preliminary study, the only target group was represented by freshwater bony fishes. The MiFish primer set (Miya et al., 2015) which amplifies a fragment of the mitochondrial *12s* rRNA gene was selected for the metabarcoding pipeline. Even if this set of primers was designed on tropical marine fishes, a large number of studies demonstrated its efficiency also in both lentic and lotic freshwater environment (McDevitt et al., 2019; Miya et al., 2020; Shu et al., 2021; Bylemans et al., 2018).

For the 2020 campaign in addition to the MiFish primer set for the fish community assessment, another set of primer was chosen to investigate also the freshwater invertebrate community after seeing the the success of the laboratory protocol and sequencing on fishes. To assess the biodiversity of this group many set of primers are available mainly targeting the cytochrome c oxidase I (*coi*) barcoding region (Leray et al., 2013; Elbrecht & Leese, 2017; Macher et al., 2018), even though alternative metabarcode primers for different genetic regions are taken into consideration (Deagle et al., 2014). In order to find the most suitable primer set for this part of the study, four different couples of primers were tested *in silico* on six taxa related to aquatic environments (Acoelomorpha, Nematoda, Clitellata, Gastropoda, Bivalvia, Arthropoda) performing simulated PCR using the *ecoPCR* program in OBITools. All the *ecoPCR* were performed on *coi* sequences downloaded from NCBI (accessed 03.02.2020). Thanks to the ecotaxstat and ecotaxspecificity commands the coverage and the resolution were tested for each primer on each taxon. The coverage is calculated as a fraction of the taxids retrieved in the database that are also in the *ecoPCR* output, while the resolution corresponds to the unambiguously identified taxa among the total number of taxa. The set of primers tested are the following: mlCOIintF + jgHCO2198 (Leray et al., 2013; Geller et al. 2013), BF3 + BF2 (Elbrecht & Leese, 2017; Elbrecht et al., 2019), FWHF2 + FWHR2n (Vamos et al., 2017), FWHF2 + EPTDr2n (Vamos et al., 2017; Leese

et al., 2021). The results of ecotaxstat and ecotaxspecificity are included in Tab. M2. Of all the primer pairs tested, FWHF2 + FWHR2n was chosen since it showed the highest coverage and good levels of specificity. BF showed the highest specificity, but the coverage was low compared to FWH.

The metabarcoding primers have a sample tag made of 8 bases with at least 3 differences out of 8 bases. In order to increase sequence variability, a variable number, ranging from 2 to 4 of fully degenerate nucleotides (leading Ns) was added.

Table M2. Results of the OBITools command ecotaxstat and ecotaxspecificity on the four different *coi* primer sets used on taxa related to aquatic environments

TAXA	PRIMER	COVERAGE		SPECIFICITY	
Acoelomorpha	mlCOIintF + jgHCO2198	Family	10.53	Family	100
		Genus	4.26	Genus	100
		Species	2.44	Species	100
	BF3 + BF2	Family	10.53	Family	100
		Genus	4.26	Genus	100
		Species	2.44	Species	100
	FWHF2 + FWHR2n	Family	100	Family	100
		Genus	85.11	Genus	100
		Species	71.95	Species	100
	FWHF2 + EPTDr2n	Family	42.11	Family	100
		Genus	21.28	Genus	100
		Species	20.73	Species	100
Nematoda	mlCOIintF + jgHCO2198	Family	34.48	Family	95.00
		Genus	28.82	Genus	97.59
		Species	21.65	Species	87.27
	BF3 + BF2	Family	36.21	Family	95.24
		Genus	29.51	Genus	97.65

		Species	23.33	Species	91.56
	FWHF2 + FWHR2n	Family	57.76	Family	97.01
		Genus	39.24	Genus	98.23
		Species	31.10	Species	81.65
	FWHF2 + EPTDr2n	Family	19.83	Family	100
		Genus	11.46	Genus	100
		Species	7.28	Species	89.19
Clitellata	mlCOIintF + jgHCO2198	Family	47.37	Family	100
		Genus	29.76	Genus	97.33
		Species	6.48	Species	96.94
	BF3 + BF2	Family	50.00	Family	100
		Genus	33.73	Genus	97.65
		Species	7.93	Species	96.67
	FWHF2 + FWHR2n	Family	100	Family	84.21
		Genus	98.02	Genus	84.62
		Species	97.29	Species	53.60
	FWHF2 + EPTDr2n	Family	92.11	Family	77.14
		Genus	94.84	Genus	81.17
		Species	95.64	Species	51.21
Gastropoda	mlCOIintF + jgHCO2198	Family	28.84	Family	95.41
		Genus	13.05	Genus	96.00
		Species	6.69	Species	85.35
	BF3 + BF2	Family	35.45	Family	94.78
		Genus	16.09	Genus	96.22
		Species	7.32	Species	91.24
	FWHF2 + FWHR2n	Family	98.15	Family	81.13
		Genus	97.65	Genus	87.04
		Species	95.57	Species	67.38
	FWHF2 + EPTDr2n	Family	83.33	Family	80.63

		Genus	70.07	Genus	84.11
		Species	55.87	Species	55.78
Bivalvia	mlCOIintF + jgHCO2198	Family	29.03	Family	100
		Genus	17.43	Genus	95.60
		Species	6.55	Species	86.19
	BF3 + BF2	Family	37.63	Family	100
		Genus	20.50	Genus	98.13
		Species	7.57	Species	90.91
	FWHF2 + FWHR2n	Family	88.17	Family	78.05
		Genus	79.12	Genus	78.21
		Species	70.71	Species	42.40
	FWHF2 + EPTDr2n	Family	34.41	Family	81.25
		Genus	31.03	Genus	73.46
		Species	25.67	Species	39.07
Arthropoda	mlCOIintF + jgHCO2198	Family	46.56	Family	95.31
		Genus	23.64	Genus	96.47
		Species	4.02	Species	81.14
	BF3 + BF2	Family	47.22	Family	95.17
		Genus	21.83	Genus	96.76
		Species	3.38	Species	87.56
	FWHF2 + FWHR2n	Family	95.25	Family	NA
		Genus	88.80	Genus	NA
		Species	84.80	Species	NA
	FWHF2 + EPTDr2n	Family	82.76	Family	NA
		Genus	73.89	Genus	NA
		Species	67.57	Species	NA

PCR

PCR amplification was performed using a single-step protocol with 10 technical replicates for each sample. The PCR mix was as follows: 10 μl AmpliTaq Gold 360 Master Mix (Applied Biosystems); 0.16 μl of BSA 20 μg/μl; 1.0 μl of each primer (5 μM); 4.84 μl ultrapure water UV treated; 3 μl of DNA template (1 ng/μl). For the MiFish primers the PCR programme consisted in a first step at 95°C for 10 min, needed to denature the Taq polymerase blocking antibody, a second step repeated for 40 cycles at 95°C for 30 sec, 60°C for 45 sec and 72°C for 30 sec, and the final extension at 72°C for 5 min. For the FWH primers the PCR settings consisted in a first step at 95°C for 10 min, a second step repeated for 40 cycles at 95°C for 30 sec, 55°C for 45 sec and 72°C for 30 sec, and the final extension at 72°C for 5 min.

Eight positive controls, eight negative controls (PCR water as template), eight extraction blanks and 12 blanks (no primer, no template) were included in each PCRs. Positive controls for FWH marker consisted in genomic material extracted from *Notospermus geniculatus* (Nemertea), *Patella rustica* (Mollusca Gastropoda), *Muscolista senhousia* (Mollusca Bivalvia), *Sarax ioanniticus* (Arthropoda Aracnida), *Glycera dibranchiata* (Arthropoda Myriapoda), *Pleoticus muelleri* (Crustacea), *Trychopeplus lancinatus* (Arthropoda Hexapoda); for MiFish marker they included DNA extracted from *Acanthobrama persidis, Acanthobrama microlepis, Chondrostoma regium, Rutilus frisii, Leuciscus vorax*, all cyprinid fishes from Middle-eastern Europe.

Amplicon lengths was 163-185 bp for the MiFish marker and 205 bp for the FWH marker.

Electrophoresis

Once amplified all the samples, the success of amplifications was checked through gel electrophoresis in a 1.5 % agarose gel stained with GelRed (Cambridge Bioscience).

Pooling

All samples from one reaction plate were pooled in a single Eppendorf and purified using MinElute columns to remove DNA fragments below 70 bp and to concentrate DNA around 10 times.

Each purified plate was than quantified by the Qubit fluorometer applying the BR assay. All the purified plates from each marker were pooled together so that the concentration of the final pool for each metabarcoding marker was 1 μg in 30 μl of DNase-free water.

Library preparation and sequencing

The PCR products were then sent to Fasteris SA (https://www.fasteris.com/en-us) for the library preparation and the Illumina sequencing.

The MiFish pool from the preliminary study consisted in 12 biological samples with 6 blanks, with 25 replicates each.

The MiFish pool from the 2020 campaign included 12 biological samples, with 12 extraction replicates and 5 PCR replicates each. This pool also included 15 samples from a non-related project targeting sea-water fish species. The Fwh pool included 12 biological samples, with 12 extraction replicates and 10 PCR replicates each. Illumina libraries were built using the MetaFast protocol provided by Fasteris SA. The two distinct libraries were run in a single Illumina MiSeq run using two different flow cells with 2 x 150 bp v2 chemistry with 5% PhiX spike-in.

3.2.6 Bioinformatic Analysis

The reads from the preliminary study were analyzed through a bioinformatic pipeline by Leray & Knowlton (2017) which relies on the OBITools metabarcoding programs package (Boyer at al. 2016). Since the main concern was the verification of the sequencing success, for the preliminary test only the OBITools pipeline was adopted.

On the other hand,to check for differences between clustering and denoising steps, the reads from the 2020 campaign were also analyzed using a QIIME2 pipeline. The OBITools pipeline was chosen since it can perform a clustering step reconstructing MOTUs, while QIIME2 provides a denoising step retriving ASVs.

As far as the OBITools pipeline is concerned, the following steps were performed:

- the alignment of paired-ends was performed using *illuminapairedend*, only reads with a quality score higher than 40.00 were kept.
- Demultiplexing and primers trimming were executed with *ngsfilter*.
- Primer-dimers, non-specific amplifications originated from library preparation artefacts and reads with ambiguous bases were removed applying a length filter selecting sequences of 140-190 bp for MiFish primers and 195-215 bp for FWH using the *obigrep* program.
- Dereplication - the clustering of highly identical sequences - was performed using *obiuniq*.
- Chimera removal on a sample-by-sample basis was applied using VSEARCH(Rognes et al., 2016) running the uchime *de novo* algorithm (Edgar et al., 2011).
- The clustering step, that is the definition of molecular operational taxonomic units (MOTUs), was performed using SWARM a step-by-step aggregation algorithm. (Sales et al., 2018) with a distance value of d=3 for MiFish and d=13 for FWH.
- After the deletion of singleton MOTUs, the taxonomic assignment was carried out using *ecotag*. For fishes the reference database was built using the *ecoPCR* program on the vertebrate standard sequences of the EMBL database, whereas for the metazoan the *ecoPCR* was performed on the invertebrate standard sequences of EMBL.

After the taxonomic assignment all the MOTUs that do not account for a minimum of 0.01% of total abundance in at least one sample and with a similarity lower than 98% were removed.

Non-target species, such as mammals, reptiles, birds ect., and those from probable carry-over contaminations were excluded (Li et al., 2018; Ushio et al., 2018).

Regarding the QIIME2 pipeline the steps were:

- first of all, raw reads were demultiplexed and primer sequences trimmed using *cutadapt*.
- The sequences were then imported and denoised using the *qiime dada2 denoise-paired* command. DADA2 allows to denoise the reads, join paired-end sequences, and also remove any chimeric sequences.

- The taxonomic assignation was performed using the *qiime feature-classifier classify-consensus-blast* command, the FeatureData[Taxonomy] object was built retrieving the taxonomic feature in the correct format for the database with the *taxonomizr* R package. The same reference databases originated for the OBITools workflow were used.

Sequences not present at least in 2 different samples and less than 10 times represented were discarded as well as all the non-target species.

3.3 RESULTS

3.3.1 Traditional survey

Fish

The traditional fish survey, as already reported, was possible only during the preliminary 2019 study and summer and fall 2020 campaigns owing to too low water level in winter and to covid-19 safety protocol in spring.

In the 2019 preliminary campaign only four species were recognized. (Table R1). The most abundant were *Cyprinus carpio* (27 individuals) and *Carassius sp*. (23 individuals); the first one was not found at site 1, whereas the second one was not present at site 3. *Pseudorasbora parva* was present at all sites with a total of 9 individuals. At site 2 only one individual of *Ameiurus sp.* was found.

Table R1. Number of individuals collected per site using the electrofishing technique during summer 2019

Species	**Site 1**	**Site 2**	**Site 3**
Carassius sp.	0	9	14
Cyprinus carpio	22	5	0
Pseudorasbora parva	4	3	2
Ameiurus sp.	0	1	0

During the summer 2020 sampling (Tab. R2), *Cyprinus carpio* (10 individuals), *Carassius sp*. (one specimen, site 2) and *Pseudorasbora parva* (23 specimens) were also found. New recorded taxa were *Gambusia holbrooki* (18 individuals), *Alburnus alborella* at site 1 (one individual) and *Lepomis gibbosus* at site 3 (one individual).

During the fall 2020 sampling four species were recognized (Tab. R2). *Pseudorasbora parva* (81 individuals) was the most abundant and present at all sites. With seven individuals each also

Ameiurus melas and *Cyprinus carpio* were present at all sites. *Gambusia holbrooki* was present at site 1 and 2.

Throughout the course of all the sampling a numerous amount of *Procambarus clarkii* in all the sampling sites was also found using the electrofishing technique.

Table R2. Number of fishes collected per site using the electrofishing technique during summer and autumn 2020

	Summer 2020			**Autumn 2020**		
Species	**Site 1**	**Site 2**	**Site 3**	**Site 1**	**Site 2**	**Site 3**
Alburnus alborella	1	0	0	0	0	0
Ameiurus melas	0	0	0	4	1	2
Carassius sp.	0	1	0	0	0	0
Cyprinus carpio	3	5	2	1	2	4
Gambusia holbrooki	3	12	3	6	1	0
Lepomis gibbosus	0	0	1	0	0	0
Pseudorasbora parva	11	5	7	54	11	16

Invertebrates

Invertebrate sampling led to 13 specimens collected with the entomological net and 23 individuals caught with the surber sampler. All specimens (either pinned, the former, or in 70 % ethanol, the latter) were identified in order to determine the most inclusive taxon level; through the morphological dichotomic keys available it was in fact possible to identify some individuals at the species level while some only at the order or family level. The results are shown in Tab. R3.

The specimens sampled using the net included: one specimen each for the orders Orthoptera, Homoptera, Coleoptera, Diptera and Lepidoptera, two Odonata, and five Heteroptera; one specimen of *Procambarus clarkii* was captured on the gutter bank near site 1.

The specimens collected with the surber sampler comprised three Gastropoda and 20 Insecta. Most of the water insects were larvae or nymphs. The only adults were four individuals of

Corixinae all belonging to the same morphospecies. A total of eight Odonata nymphs were caught: they were all zygopterans with the exception of *Anax imperator* which is an anisopteran. Two nymphs of Ephemeroptera were also collected, both identified as member of the Baetidae family, and four larvae of Chironomidae.

Table R3. List of the specimens identified using the morphological and barcoding approaches (id numbers refer to samples collected with the entomological net while letters with the surber sampler; % refers to the similarity scores of the first match (Top %) and the 100th (Low %). Photographs of the individuals in the Supplementary Materials (Fig. S8-S38).

id	**Morphological identification**	**Barcoding identification on BOLD**	Top %	Low %
1	*Platycnemis pennipes*	*Platycnemis pennipes*	99.66	85.64
2	*Platycnemis pennipes*	*Platycnemis pennipes*	100,00	85.37
3	Rynchota - Heteroptera	-		
4	Rynchota - Heteroptera	*Notostira elongata*	99.67	87.81
5	Rynchota - Heteroptera	*Notostira elongata*	100,00	87.37
6	Pipunculidae	-		
7	Piesmidae	-		
8	Orthoptera	*Chorthippus dorsatus*	99.35	94.14
9	Rynchota - Homoptera	-		
10	Rynchota	*Trigonotylus caelestialium*	100,00	88.55
11	Coleoptera	*Oedemera flavipes*	100,00	87.3
12	Geometridae	*Peribatodes rhomboidaria*	100,00	98.56
13	*Procambarus clarkii*	-		
a	*Coenagrion* sp.	-		
b	*Coenagrion* sp.	-		
c	*Valvata piscinalis*	-		
d	*Procleon bifidum*	-		
e	*Physa fontinalis*	-		
f1	Corixinae	-		

f2	*Sigara* sp.	*Sigara lateralis*	99.83	90.05
f3	Corixinae	-		
f4	Corixinae	-		
f5	Corixinae	*Sigara lateralis*	100,00	90.05
g	Tr. Chironomini	*Cryptochironomus obreptans*	100,00	96.37
h	Chironomidae	*Parachironomus gracilior*	99.67	87.87
i	*Procleon bifidum*	*Cloeon dipterum*	100,00	92.61
j	*Coenagrion* sp.	-		
k	*Anax imperator*	*Anax* sp. 1	100,00	98.93
l	Orthocladinae	-		
m	*Coenagrion* sp.	*Ischnura elegans*	99.85	99.39
n	*Coenagrion* sp.	*Ischnura elegans*	100,00	99.51
o	*Coenagrion* sp.	*Ischnura elegans*	99.84	99.26
p	*Coenagrion* sp.	*Ischnura elegans*	100,00	99.69
q	*Coenagrion* sp.	*Ischnura pumilio*	100,00	89.52
r	Valvata	-		
z	Chironomidae (eggs)	-		

Barcoding

Since the morphological identification of invertebrates was performed by a non-specialist, a molecular approach was also implemented as data support and comparison between methods.

Only 19 sequences were obtained, because for six samples the PCR was not successful (samples 6, c, d, e, r, z) and for 11 samples the sequencing failed (samples 3, 7, 9, 13, a, b, f1, f3, f4, j, l). All the BOLD top percentage for the taxonomical assignment were greater than 99.35% (Tab. R3) The morphological identification of *Anax imperator* and the two specimens of *Platycnemis pennipes* matched the molecular taxonomical assignation. On the other hand, *Procleon bifidum* was identified as *Cleon dipterum* and four *Coenagrion* sp. as *Ischnura elegans* and one specimen

as *Ischnura pumilio.* Nine samples identified at the family (samples f5, g, h, 12) or the order (samples 4, 5, 8, 10, 11) levels were assigned to species of the corresponding family/order. The sample f2 identified at genus level with the morphological approach, was assigned to a corresponding species at the genus level with barcoding.

3.3.2 Metabarcoding

Here the results of the various metabarcoding pipelines are presented. The OBITools pipeline was adopted for the analysis of the MiFish marker from the preliminary study and the 2020 campaign and the FWH marker from the 2020 campaign, while the QIIME2 pipeline was implemented for the analysis of the MiFish and FWH markers from the 2020 campaign only. The taxonomy retrieved and present in all the tables is based on the NCBI taxonomy. The number of reads illustrated in every heatmap is normalized for a better graphical comparison.

Mifish

OBITools pipeline

The metabarcoding analysis of the preliminary sampling (summer 2019) gave a total of 18 845 197 raw reads with 5987857 sequences finally retrieved following the OBITools pipeline (Tab. R4).

Table R4. OBITools output on the MiFish sequences from the preliminary test (summer 2019). MOTUs assignation, Percentage of identity, Order, Family, reads per site per different filtration method and total number of reads are shown.

				Prefiltering in the field			Without prefiltering			
Species	**Best Identity**	**Order**	**Family**	**Site 1**	**Site 2**	**Site 3**	**Site 1**	**Site 2**	**Site 3**	**Total Reads**
Abramis brama	100%	Cypriniformes	Leuciscidae	0	0	0	22804	0	0	22804
Alburnus alburnus	100%	Cypriniformes	Leuciscidae	0	0	0	1	10677	0	10678
Alburnus arborella	100%	Cypriniformes	Leuciscidae	0	0	10037	1	0	5909	15947
Ameiurus melas	100%	Siluriformes	Ictaluridae	0	0	0	5837	0	290	6127
Carassius sp.	99%	Cypriniformes	Cyprinidae	0	0	10617	12583	13175	85540	121915
Cyprininae	99%	Cypriniformes	Cyprinidae	27026	48	112480	2551084	1544223	615452	4850313
Dicentrarchus labrax	99%	Perciformes	Moronidae	0	0	0	0	0	422	422
Gambusia holbrooki	100%	Cyprinodontiformes	Poeciliidae	142	0	1	113555	6214	1861	121773
Gobio gobio	100%	Cypriniformes	Gobionidae	0	0	10005	0	0	1	10006
Pseudorasbora sp.	100%	Cypriniformes	Gobionidae	11597	6	59597	314246	252932	189494	827872

A total of 10 MOTUs were detected, belonging to 10 different taxa.

Carassius sp., *Pseudorasbora* sp. and *Ameiurus melas* were found also during the electrofishing survey, moreover in the eDNA species list a Cyprininae species was detected at all sites. In the metabarcoding analysis *Abramis brama* was found only at site 2, *Gobio gobio* only at site 3. *Alburnus alburnus* was present at sites 1 and 2, while *Alburnus arborella* at sites 1 and 3. *Gambusia holbrooki* was detected in all sampling sites. *Dicentrachus labrax,* a common sea or brackish fish, was found only in one sample (Fig. R1).

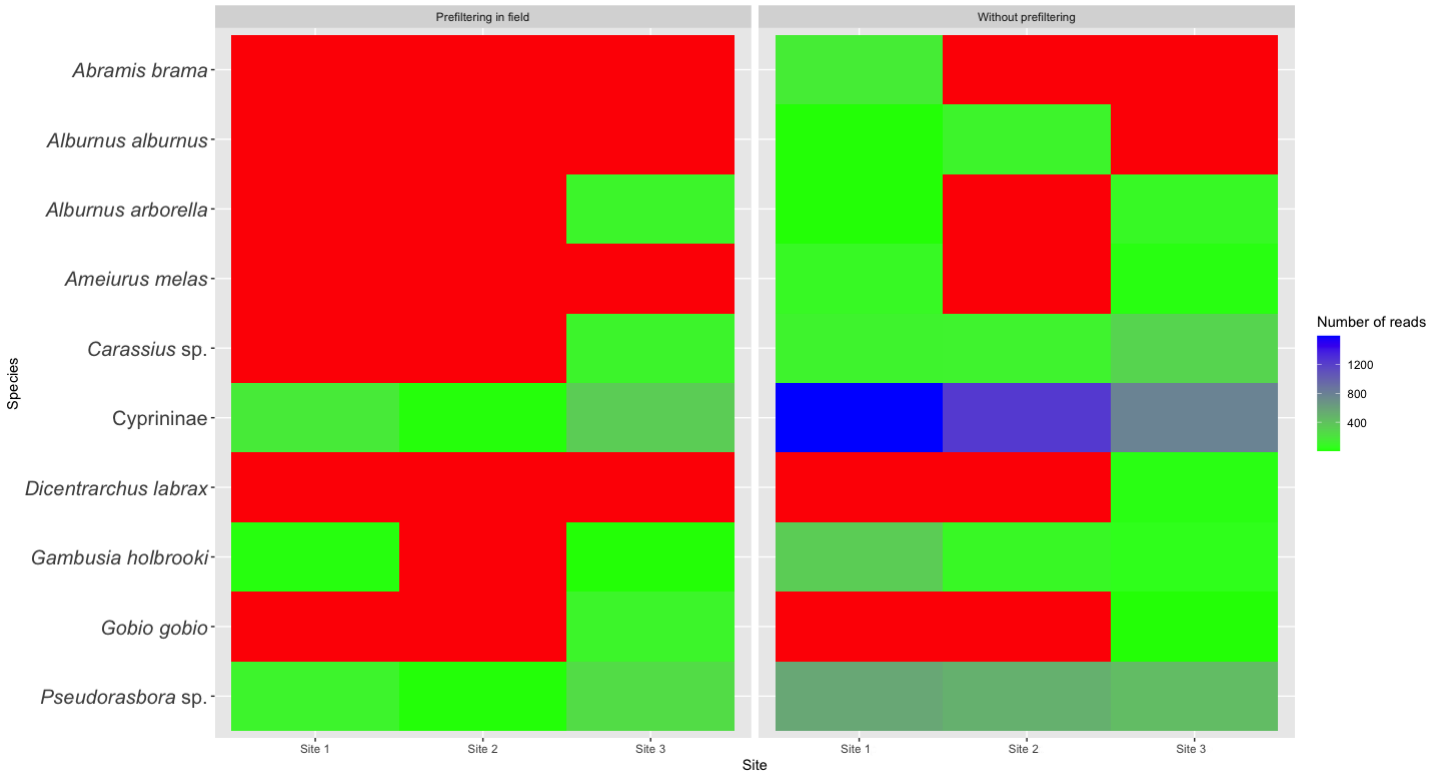


Figure R1. Heatmap showing the number (normalized) of MiFish reads of each OBITools MOTUs per site for the 2019 summer sampling. Red color means no read assigned to that sample.

18448274 raw reads were obtained for the MiFish Illumina sequencing of the samples belonging to the 2020 campaigns (including the sequences from the side project carried out with Hydrosyngery co-op in Grado as mentioned in chapter 1). After all the filtering steps 2 121 835 reads were retrieved. Raw number of reads per sample and are shown in Tab. R5.

Table R5. OBITools species list on the MiFish sequences from the 2020 campaign. MOTUs assignation, Percentage of identity, Order, Family, reads per site per different seasonal sampling and total number of reads are shown.

				Winter			Spring			Summer			Autumn			
Species	**Percentage of id**	**Order**	**Family**	**Site 1**	**Site 2**	**Site 3**	**Site 1**	**Site 2**	**Site 3**	**Site 1**	**Site 2**	**Site 3**	**Site 1**	**Site 2**	**Site 3**	**Total Reads**
Ameiurus nebulosus	99%	Siluriformes	Ictaluridae	2150	18879	12968	7	0	0	0	0	3	51695	16083	16170	117955
Carassius sp.	100%	Cypriniformes	Cyprinidae	22682	233560	22286	42108	2730	33	13	3	0	6238	14206	18257	362116
Cyprinus carpio	100%	Cypriniformes	Cyprinidae	167651	113087	32170	59965	30478	66895	22373	36065	64498	17793	22505	15124	648604
Gambusia affinis	99%	Cyprinodontiformes	Poeciliidae	7510	12658	4471	51	0	1	0	3	9	76964	85099	119843	306609
Leuciscinae	98%	Cypriniformes	Leuciscidae	64	0	0	0	0	0	0	0	0	0	0	0	64
Leuciscus leuciscus	99%	Cypriniformes	Leuciscidae	116234	1193	496	2	151	9618	721	3580	5	19	0	12	132031
Pseudorasbora parva	98%	Cypriniformes	Gobionidae	41707	41648	29963	31895	4272	5098	12	5008	8	95806	132776	150340	538533

A total of seven MOTUs were identified, corresponding to 7 different taxa of fishes.

Carassius sp., *Cyprinus carpio*, and *Pseudorasbora parva* were also found in the traditional survey. In the eDNA species list *Ameiurus nebulosus* was detected while during the electrofishing session *Ameiurus melas* was identified. The same applies also to *Gambusia* specimens*:* in fact in the metabarcoding analysis it is present as *Gambusia affinis* whereas in the traditional survey various *Gambusia holbrooki* were identified in the field. *Leuciscus leuciscus* and another Leuciscinae species are found only in the metabarcoding analyses.

Pseudorasbora parva and *Cyprinus carpio* were detected in all samplings at all sites (Fig. R2).

In the winter 2020 survey all species were found at all sites, except for the Leuciscinae taxon which was found only at site 1.

In the spring 2020 sampling, only the Leuciscinae species was absent, while *Ameiurus nebulosus* was found only at site 1 and *Gambusia affinis* only at sites 1 and 3.

In the summer 2020 collection, the MOTU corresponding to Leuciscinae was absent. *Ameiurus nebulosus* was found only at site 3, *Carassius* sp. at sites 1 and 2, *Gambusia affinis* at sites 2 and 3.

In the autumn 2020 survey, molecular data do not retrieve the Leuciscinae species and indicate the presence of *Leuciscus leuciscus* only at sites 1 and 3.

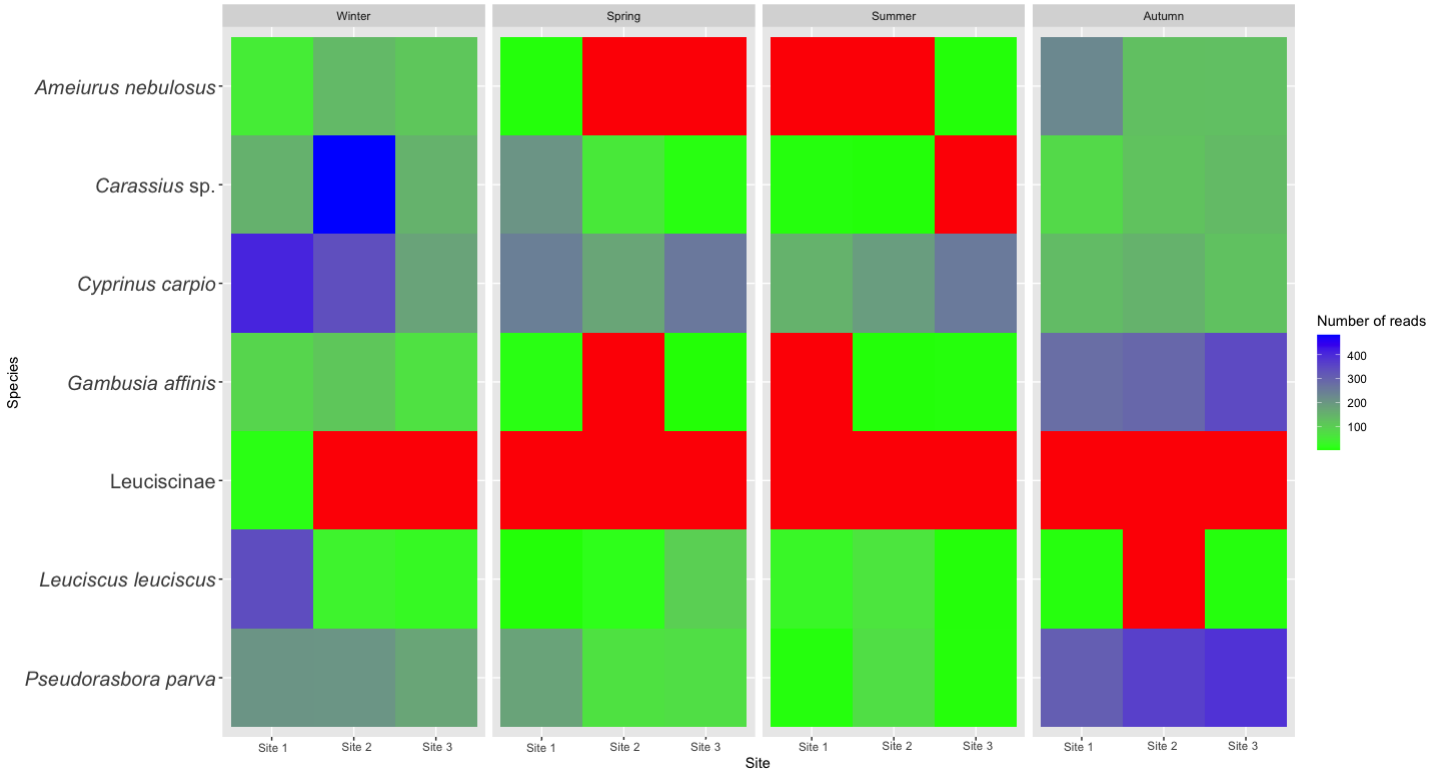


Figure R2. Heatmap showing the number (normalized) of MiFish reads of each OBITools MOTUs per site for each season of the 2020 campaign. Red color means no reads assigned to that sample

QIIME2 pipeline

This pipeline was adopted only for data from the 2020 campaign.

At the end of the QIIME2 pipeline on the MiFish reads, 2875 sequences were obtained after all the filtering steps. A total of eight amplicon sequence variants (ASVs) were retrieved belonging to an equal number of species (Tab. R6).

Tab. R6. QIIME2 output on the MiFish sequences from the 2020 campaign. ASVs assignation, Order, Family, reads per site per different seasonal sampling and total number of reads are shown.

			Winter			Spring			Summer			Autumn			
Species	**Order**	**Family**	**Site 1**	**Site 2**	**Site 3**	**Site 1**	**Site 2**	**Site 3**	**Site 1**	**Site 2**	**Site 3**	**Site 1**	**Site 2**	**Site 3**	**Total reads**
Ameiurus sp.	Siluriformes	Ictaluridae	0	0	0	25	0	12	5	0	158	2	0	0	202
Carassius gibelio	Cypriniformes	Cyprinidae	0	0	13	0	208	0	0	0	0	0	17	0	238
Carassius sp.	Cypriniformes	Cyprinidae	93	12	22	9	81	59	67	0	0	17	0	42	402
Esox lucius	Esociformes	Esocidae	0	0	13	0	0	0	0	0	0	0	0	191	204
Gambusia sp.	Cyprinodontiformes	Poeciliidae	9	10	4	24	12	19	47	20	99	0	0	13	257
Oncorhynchus mykiss	Salmoniformes	Salmonidae	16	11	10	0	234	295	6	0	82	49	0	3	706
Pseudorasbora parva	Cypriniformes	Gobionidae	26	0	23	45	50	54	77	0	80	0	0	21	376
Rutilus sp.	Cypriniformes	Leuciscidae	39	0	7	53	38	71	10	0	80	41	0	151	490

Carassius sp. and *Pseudorasbora parva* were present also in the electrofishing survey; even though during the traditional campaign *Ameiurus melas and Gambusia affinis* were identified at species level the QIIME2 pipeline assigned two ASVs only at the genera level (*Ameiurus* and *Gambusia).* Besides the ASVs assigned to the *Carassius* genus, another ASV was assigned to *Carassius gibelio.*

Esox lucius, Oncorhynchus mykiss and *Rutilus* sp. were absent both in the traditional investigation and in the OBITools pipeline.

During the winter 2020 survey, *Ameiurus* sp. was not found while all the other species were detected at all sites, with the exception of *Carassius gibelio* and *Esox lucius* found only at site 3 and *Pseudorasbora parva* and *Rutilus* sp. found at sites 1 and 3 (Fig. R3).

During the spring 2020 campaign, *Esox lucius* was absent whereas *Ameiurus* sp. was present at sites 1 and 3, *Carassius gibelio* at site 2 and *Oncorhynchus mykiss* at sites 2 and 3.

During the summer 2020 campaign, both *Carassius gibelio* and *Esox lucius* were absent and only *Gambusia* sp. was detected at all sites. All the other species were present at sites 1 and 3, with *Carassius* sp. present only at site 1.

During the autumn 2020 survey, all the species were detected but not in all the sampling sites. *Ameiurus* sp. was present at site 1, while *Carassius gibelio* at site 2. All the other species were found at site 3, *Carassius* sp., *Oncorhynchus mykiss* and *Rutilus* sp. also at site 1.

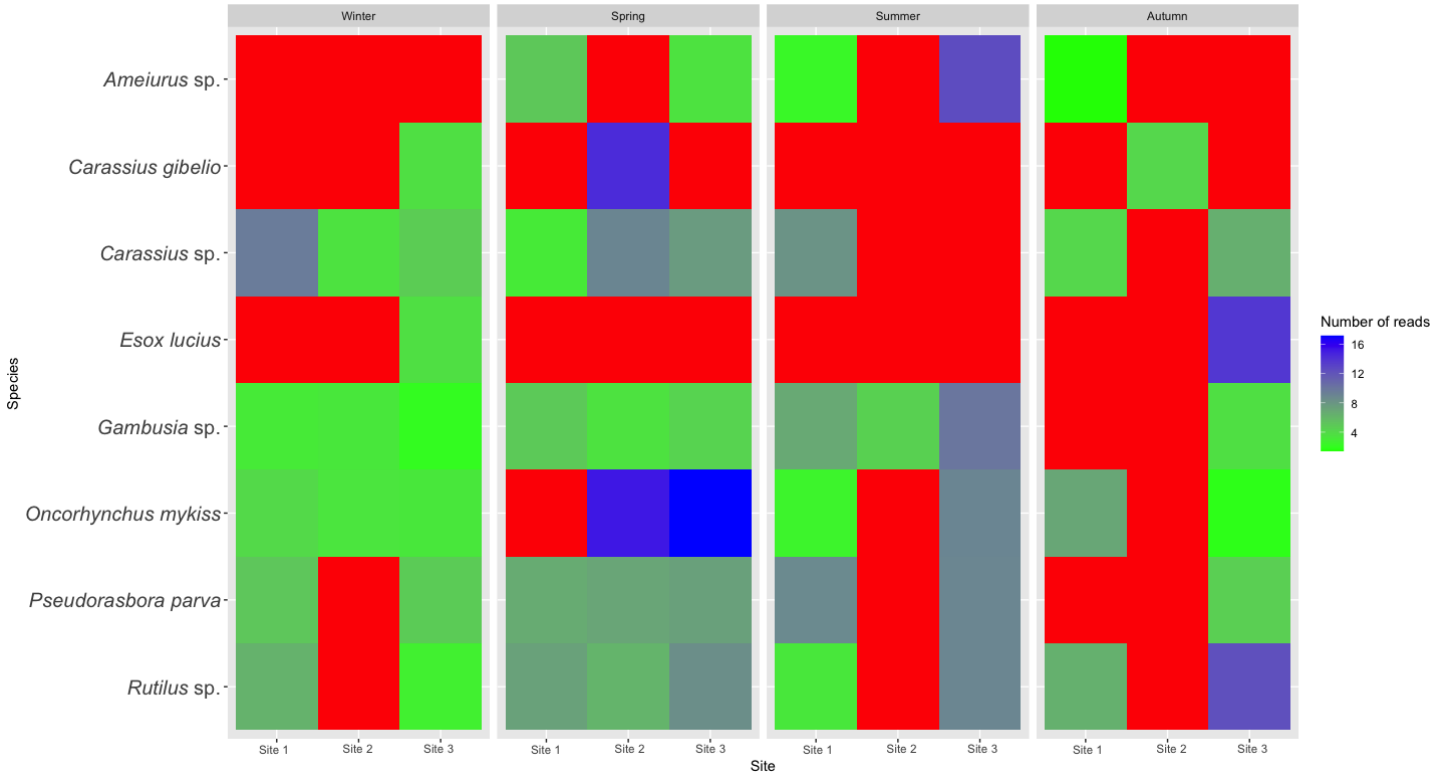


Figure R3. Heatmap showing the number (normalized) of MiFish reads of each QIIME2 ASVs per site for each season. Red color means no reads assigned to that sample.

Fwh

OBITools pipeline

The sequencing of the FWH library resulted in a total of 9927434 raw reads, after filtering 1246200 metazoan reads were retrieved. Further reads information are listed in Table R7.

Table R7. OBITools species list on the FWH sequences from the 2020 campaign. MOTUs assignation, percentage of identity, Phylum, Class, Order, Family, reads per site per different seasonal sampling and total number of reads are shown.

						Winter			Spring			Summer			Autumn			
Species	**Percentage of id.**	**Phylum**	**Class**	**Order**	**Family**	**Site 1**	**Site 2**	**Site 3**	**Site 1**	**Site 2**	**Site 3**	**Site 1**	**Site 2**	**Site 3**	**Site 1**	**Site 2**	**Site 3**	**Total Reads**
Acanthocyclops americanus	100%	Arthropoda	Maxillopoda	Cyclopoida	Cyclopidae	2	18	5	97	429	534	5904	1679	18876	19	711	6	28280
Aculops lycopersici	100%	Arthropoda	Arachnida	Trombidiformes	Eriophyidae	1	0	1	0	0	1	0	0	0	0	0	92	95
Amphichaeta raptisae	100%	Annelida	Oligochaeta	Haplotaxida	Naididae	0	83	0	0	0	0	0	0	0	0	0	0	83
Anuraeopsis sp.	100%	Rotifera	Monogononta	Ploima	Brachionidae	3	3	4	0	0	1	2	6	4	3089	4712	452	8276
Asplanchna sieboldi	100%	Rotifera	Monogononta	Ploima	Asplanchnidae	9	50	2	3	0	3	0	0	0	20	15	0	102
Balclutha frontalis	100%	Arthropoda	Insecta	Hemiptera	Cicadellidae	0	0	0	0	0	0	0	0	0	55	1	0	56
Brachionus angularis	100%	Rotifera	Monogononta	Ploima	Brachionidae	1	0	0	1	0	375	0	0	0	4	12	21	414
Brachionus quadridentatus	100%	Rotifera	Monogononta	Ploima	Brachionidae	0	2	3	3	2	517	269	354	1294	307	2	73	2826
Brachionus variabilis	100%	Rotifera	Monogononta	Ploima	Brachionidae	0	0	0	20	1	81	0	0	0	18	0	178	298
Branchiura sowerbyi	100%	Annelida	Oligochaeta	Haplotaxida	Naididae	0	12	7	0	0	0	0	0	0	2	15	17	53
Caenorhabditis brenneri	99%	Nematoda	Chromadorea	Rhabditida	Rhabditidae	749	894	0	2	0	1	0	0	0	0	2	478	2126
Chaetogaster diastrophus	99%	Annelida	Oligochaeta	Haplotaxida	Naididae	27	60	189	0	0	0	0	0	0	0	2	12	290
Chironomus sp.	99%	Arthropoda	Insecta	Diptera	Chironomidae	0	0	1	1	4	44	0	0	0	0	17	40	107
Chironomus transvaalensis	99%	Arthropoda	Insecta	Diptera	Chironomidae	0	38	35	2	0	1	0	0	0	77	168	103	424
Cricotopus sp.	100%	Arthropoda	Insecta	Diptera	Chironomidae	0	0	0	21	65	3	0	0	0	0	0	0	89
Culicidae	100%	Arthropoda	Insecta	Diptera	Culicidae	0	0	0	0	0	0	0	0	0	22	0	0	22
Daphnia magna	100%	Arthropoda	Branchiopoda	Diplostraca	Daphniidae	0	0	0	35	0	12	0	0	0	0	0	0	47
Dero digitata	100%	Annelida	Oligochaeta	Haplotaxida	Naididae	0	2	9	0	0	0	0	0	0	69	76	60	216
Deroceras sp.	100%	Mollusca	Gastropoda	Stylommatophora	Agriolimacidae	102	4	0	0	0	0	1	0	0	20	0	0	127
Dicrotendipes nervosus	99%	Arthropoda	Insecta	Diptera	Chironomidae	0	0	0	0	0	0	0	0	0	0	0	94	94
Eucyclops cf. *estherae*	100%	Arthropoda	Maxillopoda	Cyclopoida	Cyclopidae	0	0	25	0	0	0	0	0	0	0	0	0	25
Formica exsecta	100%	Arthropoda	Insecta	Hymenoptera	Formicidae	18	1	0	0	0	0	0	0	0	119	306	177	621
Globulidrilus riparius	99%	Annelida	Oligochaeta	Enchytraeida	Enchytraeidae	0	0	0	0	0	0	0	0	0	0	0	108	108
Glyptotendipes sp.	99%	Arthropoda	Insecta	Diptera	Chironomidae	38	0	0	0	0	0	0	0	0	0	0	0	38
Hydra vulgaris 1	100%	Cnidaria	Hydrozoa	Anthoathecata	Hydridae	0	29	63	0	0	0	0	0	0	1	49	0	142
Hydra vulgaris 2	100%	Cnidaria	Hydrozoa	Anthoathecata	Hydridae	0	0	0	0	0	0	0	0	2	0	0	17	19
Hypogastrura vernalis	100%	Arthropoda	Hexapoda	Poduromorpha	Hypogastruridae	0	0	0	0	0	0	0	0	0	91	60	0	151
Hypoponera opacior	99%	Arthropoda	Insecta	Hymenoptera	Formicidae	0	0	0	0	0	0	0	0	0	123	0	0	123
Keratella cochlearis 1	100%	Rotifera	Monogononta	Ploima	Brachionidae	0	0	0	0	0	0	0	0	0	16	0	0	16
Keratella cochlearis 2	99%	Rotifera	Monogononta	Ploima	Brachionidae	25	30	0	8	616	611	1	0	1	8	0	0	1300
Keratella sp.	100%	Rotifera	Monogononta	Ploima	Brachionidae	0	56	0	0	0	0	0	0	0	1	21	0	78
Limnodrilus hoffmeisteri 1	100%	Annelida	Oligochaeta	Haplotaxida	Naididae	17	4	0	0	0	0	0	0	0	0	0	249	270
Limnodrilus hoffmeisteri 2	100%	Annelida	Oligochaeta	Haplotaxida	Naididae	0	5	48	0	0	0	0	0	0	46	55	0	154
Limnodrilus sp.	100%	Annelida	Oligochaeta	Haplotaxida	Naididae	226	107	27	1	3	0	0	0	0	47	55	55	521
Liposcelis rufa	100%	Arthropoda	Insecta	Psocoptera	Liposcelididae	0	1	1	0	0	0	0	0	0	1	1783	0	1786
Macrothrix sp.	100%	Arthropoda	Branchiopoda	Diplostraca	Macrotrichidae	1	0	0	18	18	38	35	5	18	0	0	0	133
Myrmica sp.	100%	Arthropoda	Insecta	Hymenoptera	Formicidae	0	0	1	0	0	0	0	0	0	7	49	10	67
Nais barbata	100%	Annelida	Oligochaeta	Haplotaxida	Naididae	55	53	186	0	0	0	0	0	0	0	0	0	294
Nais christinae	100%	Annelida	Oligochaeta	Haplotaxida	Naididae	11	117	0	0	0	0	0	0	0	0	0	0	128
Nais communis/variabilis complex sp. 1	100%	Annelida	Oligochaeta	Haplotaxida	Naididae	308	249	208	1	1	2	1	0	0	0	108	1	879
Nais communis/variabilis complex sp. 2	99%	Annelida	Oligochaeta	Haplotaxida	Naididae	0	44	3	0	0	0	0	0	0	0	0	0	47
Nylanderia sp.	100%	Arthropoda	Insecta	Hymenoptera	Formicidae	2	0	0	6	2	0	0	0	1	0	0	0	11
Othius laeviusculus	100%	Arthropoda	Insecta	Coleoptera	Staphylinidae	0	0	0	0	0	1	0	0	1	76	0	0	78
Paranais frici	100%	Annelida	Oligochaeta	Haplotaxida	Naididae	44	126	217	0	0	0	0	0	0	1	0	1	389
Phragmatobia fuliginosa	100%	Arthropoda	Insecta	Lepidoptera	Erebidae	0	0	0	0	0	0	0	0	0	0	0	78	78
Physella acuta	100%	Mollusca	Gastropoda	Basommatophora	Physidae	3141	4055	3322	9	7	10	9	7	4	816	968	1101	13449
Polyarthra dolichoptera	100%	Rotifera	Monogononta	Ploima	Synchaetidae	1	1	2	4	0	1	117	683	10315	3	22	7	11156
Polyarthra dolichoptera complex sp.	99%	Rotifera	Monogononta	Ploima	Synchaetidae	1646	4588	7	5	2	1	1	0	0	2	2	4	6258
Polyarthra sp.	99%	Rotifera	Monogononta	Ploima	Synchaetidae	0	0	0	0	0	0	0	7	0	0	0	0	7
Procambarus clarkii	100%	Arthropoda	Malacostraca	Decapoda	Cambaridae	0	0	0	0	0	0	0	0	0	27	0	0	27
Psammoryctides barbatus	99%	Annelida	Oligochaeta	Haplotaxida	Naididae	14	0	10	0	0	0	0	0	0	25	8	0	57
Scatopsciara vitripennis	99%	Arthropoda	Insecta	Diptera	Sciaridae	0	250	0	0	0	0	0	0	0	0	0	0	250
Solenopsis fugax	100%	Arthropoda	Insecta	Hymenoptera	Formicidae	0	0	0	0	0	0	0	0	0	41	132	2	175
Stylodrilus lemani	98%	Annelida	Oligochaeta	Lumbriculida	Lumbriculidae	60	0	0	0	0	0	0	0	0	0	0	0	60
Sulcospira paludiformis	100%	Mollusca	Gastropoda	Caenogastropoda crown group	Pachychilidae	70	18	9	0	0	0	0	0	11	48	116	67	339
Tabanus autumnalis	100%	Arthropoda	Insecta	Diptera	Tabanidae	0	0	0	0	0	0	0	0	0	30	0	0	30
Thermocyclops cf. *taihokuensis*	99%	Arthropoda	Maxillopoda	Cyclopoida	Cyclopidae	68	0	0	0	0	0	0	0	0	0	15	4	87
Tubifex tubifex 1	100%	Annelida	Oligochaeta	Haplotaxida	Naididae	0	4	1	0	0	0	0	0	0	0	0	53	58
Tubifex tubifex 2	99%	Annelida	Oligochaeta	Haplotaxida	Naididae	128	7	40	0	0	0	0	0	0	0	4	1	180
Xiphinema brevicollum	100%	Nematoda	Enoplea	Dorylaimida	Longidoridae	40	0	0	0	0	0	0	0	0	0	0	0	40

A total of 60 MOTUs were defined corresponding to 54 distinct invertebrate species.

Species belonging to six phyla were retrieved (listed in alphabetical order): Annelida (14 species), Arthropoda (25 species), Cnidaria (1 species), Mollusca (3 species), Nematoda (2 species), and Rotifera (9 species).

Annelida stands for the 25,90 % of all metazoans detected, all species belong to the Haplotaxida order of the Oligochaeta class except for *Globulidrilus riparius* which is an Enchytraeida member.

Arthropoda is the most abundant phylum (46,25 %), with six crustaceans, one arachnid and 18 hexapods. The crustacean species belong to the orders Cyclopoida (three species), Diplostraca (two species) and Decapoda (one species).

Aculops lycopersici, belonging to Trombidiformes, is the only arachnid listed. Among Hexapoda, only *Hypogastrura vernalis* is a Collembola. Hemiptera, Psocoptera, Coleoptera, and Lepidoptera are represented by only one species each, while Diptera has eight representatives and Hymenoptera five.

The freshwater *Hydra vulgaris* is the only Cnidaria species found, representing the 1,85% of the total.

The three gastropods belonging to the Mollusca phylum concur to represent the 5,55 % of the invertebrates.

Caenorhabditis brenneri and *Xiphinema brevicollum* are the only members of the phylum Nematoda (3,70%).

All the rotifers belong to the Ploima order, being Rotifera the third most abundant phylum (16,65 %).

During the winter 2020 campaign, the six phyla are all present with 40 distinct species. *Eucyclops* sp.*, Amphicaeta raptisae, Glyptotendipes* sp.*, Nais barbata, Nais christinae, Scatopsciara vitripennis, Stylodrilus paludiformis and Xiphinema brevicollum* are found only in the winter survey (Fig. R4).

During the spring 2020 campaign, *Hydra vulgaris* is absent while other 21 species belonging to all the remaining phyla are represented. *Daphnia magna* and *Cricotopus* sp. are present only during this campaign.

During the summer 2020 survey, the majority of the 15 species belongs to Arthropoda and Rotifera, but also *Hydra vulgaris* (Cnidaria), *Deroceras sp.* (Mollusca), *Nais communis/variabilis* complex (Annelida), *Pysella acuta* and *Sulcospira paludiformis* (Mollusca) are present; no Nematoda are recorded. *Polyarthra* sp. (Rotifera) is found only during this seasonal survey.

All phyla were also retrieved during the autumn 2020 survey and are represented by 42 different species. *Balclutha frontalis*, one Culicidae species, *Globulidrilus riparius, Hypogastrura vernalis, Hypoponera opacior, Phragmatobia fulginosa, Solenopsis fugax, Tabanus autumnalis,* are exclusively recorded during the fall campaign.

Acanthocyclops americanus, Anuraeopsis sp.*, Brachionus quadridentatus, Keratella cochlearis, Nais communis/variabilis complex* sp., *Physella acuta, Polyarthra dolichoptera* and *Polyarthra dolichoptera complex* sp. are always detected in each seasonal sampling.

Figure R4. Heatmap showing the number (normalized) of FWH reads of each OBITools MOTUs per site for each season. Red color means no reads assigned to that sample.

QIIME2 pipeline

After the management of data through the QIIME2 pipeline, a total of 49811 reads were assigned to 45 different ASVs belonging to as many different species (Tab. R8).

Table R8. QIIME2 species list on the FWH sequences from the 2020 campaign. ASVs assignation, Phylum, Class, Order, Family, reads per site per different seasonal sampling and total number of reads are shown.

					Winter			Spring			Summer			Autumn			
Species	**Phylum**	**Class**	**Order**	**Family**	**Site 1**	**Site 2**	**Site 3**	**Site 1**	**Site 2**	**Site 3**	**Site 1**	**Site 2**	**Site 3**	**Site 1**	**Site 2**	**Site 3**	**Total reads**
Acanthocyclops americanus	Arthropoda	Hexanauplia	Cyclopoida	Cyclopidae	3	25	36	94	247	341	3250	1047	11202	53	554	16	16868
Aculops lycopersici	Arthropoda	Arachnida	Trombidiformes	Eriophyidae	0	0	0	0	0	0	0	0	0	0	0	46	46
Anuraeopsis sp.	Rotifera	Eurotatoria	Ploima	Brachionidae	3	7	0	0	0	0	15	8	7	1888	2802	269	4999
Asplanchna sieboldii	Rotifera	Eurotatoria	Ploima	Asplanchnidae	3	0	0	0	0	0	0	0	0	13	15	0	31
Brachionus angularis	Rotifera	Eurotatoria	Ploima	Brachionidae	0	0	0	0	0	56	0	0	0	0	6	17	79
Brachionus quadridentatus	Rotifera	Eurotatoria	Ploima	Brachionidae	0	9	7	0	0	440	181	311	803	265	17	55	2088
Brachionus sp.	Rotifera	Eurotatoria	Ploima	Brachionidae	0	0	0	0	16	0	0	0	0	4	0	0	20
Brachionus variabilis	Rotifera	Eurotatoria	Ploima	Brachionidae	0	0	0	0	0	59	0	0	0	9	0	144	212
Branchiura sowerbyi	Annelida	Clitellata	Haplotaxida	Naididae	0	3	6	0	0	0	0	0	0	0	8	10	27
Caenorhabditis brenneri	Nematoda	Chromadorea	Rhabditida	Rhabditidae	479	577	0	0	0	3	0	0	0	0	0	314	1373
Chaetogaster diastrophus	Annelida	Clitellata	Haplotaxida	Naididae	11	22	104	0	0	0	0	0	0	0	0	4	141
Chironomus plumosus	Arthropoda	Insecta	Diptera	Chironomidae	0	0	0	0	0	19	0	0	0	0	0	0	19
Chironomus transvaalensis	Arthropoda	Insecta	Diptera	Chironomidae	0	10	17	0	0	0	0	0	0	35	99	0	161
Cricotopus sp.	Arthropoda	Insecta	Diptera	Chironomidae	0	0	0	0	32	0	0	0	0	0	0	0	32
Daphnia magna	Arthropoda	Branchiopoda	Diplostraca	Daphniidae	0	0	0	26	0	6	0	0	0	0	0	0	32
Dero digitata	Annelida	Clitellata	Haplotaxida	Naididae	0	0	0	0	0	0	0	0	0	41	34	36	111
Dero sp.	Annelida	Clitellata	Haplotaxida	Naididae	0	0	7	0	0	0	0	0	0	4	0	0	11
Deroceras reticulatum	Mollusca	Gastropoda	Stylommatophora	Agriolimacidae	67	0	0	0	0	0	0	0	0	14	0	0	81
Diplogasteridae	Nematoda	Chromadorea	Rhabditida	Diplogasteridae	110	192	36	0	0	0	0	0	3	265	0	0	606
Eucyclops cf.	Arthropoda	Hexanauplia	Cyclopoida	Cyclopidae	0	0	16	0	0	0	0	0	0	0	0	0	16
Formica exsecta	Arthropoda	Insecta	Hymenoptera	Formicidae	0	0	0	0	0	0	0	0	0	85	24	10	119
Habrotrocha sp.	Rotifera	Eurotatoria	Philodinida	Habrotrochidae	298	112	0	0	0	0	0	0	0	0	0	0	410
Hydra vulgaris	Cnidaria	Hydrozoa	Anthoathecata	Hydridae	0	21	34	0	0	0	0	0	0	0	28	8	91
Hypogastrura vernalis	Arthropoda	Collembola	Poduromorpha	Hypogastruridae	0	0	0	0	0	0	0	0	0	41	28	0	69
Keratella cochlearis	Rotifera	Eurotatoria	Ploima	Brachionidae	0	27	0	8	495	436	0	0	6	7	0	0	979
Keratella tropica	Rotifera	Eurotatoria	Ploima	Brachionidae	0	30	0	0	0	0	0	0	0	0	11	0	41
Limnodrilus hoffmeisteri	Annelida	Clitellata	Haplotaxida	Naididae	0	0	31	0	0	0	0	0	0	27	37	0	95
Limnodrilus sp.	Annelida	Clitellata	Haplotaxida	Naididae	42	23	14	0	0	0	0	0	0	3	0	25	107
Liposcelis rufa	Arthropoda	Insecta	Psocoptera	Liposcelididae	0	0	0	0	0	0	0	0	0	0	1055	0	1055
Macrothrix sp.	Arthropoda	Branchiopoda	Diplostraca	Macrotrichidae	0	0	0	20	7	29	32	0	18	0	0	0	106
Myrmica sabuleti	Arthropoda	Insecta	Hymenoptera	Formicidae	0	0	0	0	0	0	0	0	0	0	8	6	14
Nais barbata	Annelida	Clitellata	Haplotaxida	Naididae	36	36	64	0	0	0	0	0	0	0	0	0	136
Nais communis/variabilis	Annelida	Clitellata	Haplotaxida	Naididae	164	155	102	0	0	0	0	0	0	0	56	0	477
Nais variabilis	Annelida	Clitellata	Haplotaxida	Naididae	10	65	0	0	0	0	0	0	0	0	0	0	75
Othius laeviusculus	Arthropoda	Insecta	Coleoptera	Staphylinidae	0	0	0	0	0	0	0	0	0	43	0	0	43
Paranais frici	Annelida	Clitellata	Haplotaxida	Naididae	25	83	139	0	0	0	0	0	0	0	0	0	247
Physella acuta	Mollusca	Gastropoda	NA	Physidae	1816	2245	1828	13	0	7	6	0	0	382	565	564	7426
Polyarthra dolichoptera	Rotifera	Eurotatoria	Ploima	Synchaetidae	942	2738	13	26	0	7	115	524	6417	4	113	0	10899
Polyarthra vulgaris	Rotifera	Eurotatoria	Ploima	Synchaetidae	0	19	0	0	0	0	0	0	5	0	0	0	24
Potamothrix bavaricus	Annelida	Clitellata	Haplotaxida	Naididae	0	0	6	0	0	0	0	0	0	3	17	95	121
Psammoryctides barbatus	Annelida	Clitellata	Haplotaxida	Naididae	5	0	0	0	0	0	0	0	0	11	0	0	16
Rhabditida	Nematoda	Chromadorea	Rhabditida	__	0	0	0	12	0	0	0	0	0	0	0	0	12
Solenopsis sp.	Arthropoda	Insecta	Hymenoptera	Formicidae	0	0	0	0	0	0	0	0	0	21	82	0	103
Sulcospira paludiformis	Mollusca	Gastropoda	NA	Pachychilidae	24	8	3	0	0	0	0	0	0	0	39	23	97
Tubifex tubifex	Annelida	Clitellata	Haplotaxida	Naididae	76	0	20	0	0	0	0	0	0	0	0	0	96

All the species belong to the same six phyla retrieved with the OBITools pipeline: 13 Annellida, 14 Arthropoda, one Cnidaria, three Mollusca, three Nematoda and 11 Rotifera.

The Annelida species represent the 28.89 % and are all member of the Haplotaxida order of the Oligochaeta class.

Arthropoda representatives are again the most abundant, the phylum reaching 31.11%, with four crustaceans, one arachnid and nine hexapods. Within the crustacean samples, two species belong to the Diplostraca order and two to the Cyclopoida one. *Aculops lycopersici* (Trombidiformes) is again the only arachnid species retrieved. All the hexapods belong to the Insecta class, with the exception of *Hypogastrura vernalis* which is a Collembola. Coleoptera, Psocoptera, Diptera are represented by one species only, while Diptera by two and Hymenoptera by three.

Hydra vulgaris is the only representative of the Cnidaria phylum (2.22%)

All the Mollusca members (6.67%) are gastropods and all the Nematoda (6.67%) are Rhabditida.

The Rotifera phylum is the third most abundant (24.44%) and all the species belong to the Ploima order with the exception of *Habrotrocha* sp. which belongs to the Philodinida order.

During the winter 2020 survey, all the phyla are detected with a total of 29 species. *Eucyclops* sp., *Habrotrocha* sp., *Nais barbata, Nais variabilis, Paranais frici* and *Tubifex tubifex* are found only during this season survey (Fig. R5).

The only member of the Cnidaria phylum, *Hydra vulgaris*, is absent in the winter campaign but all the remaining phyla are represented with a total of 14 species. *Cricotopus* sp., *Daphnia magna* and the Rhabdutida species are only detected during this survey.

Only nine species are detected during the summer 2020 survey, no Annellida or Cnidaria are present. Only one Nematoda, one Mollusca, two Arthropoda and five Rotifera are identified during the summer.

A total of 34 species are detected during the autumn 2020 survey all belonging to the six phyla. *Aculops lycopersici, Formica exsecta, Hypogastrura vernalis, Liposcelis rufa, Othius laeviusculus* and *Solenopsis* sp. are found exclusively during this survey.

Acanthocyclops americanus, Brachionus quadridentatus, Keratella cochlearis, Physella acuta and *Polyarthra dolichoptera* are present in every seasonal sampling.

Figure R5. Heatmap showing the number (normalized) of FWH reads of each QIIME2 ASVs per site for each season. Red color means no reads assigned to that sample.

On the whole, during winter and autumn campaigns, the eDNA results from both pipelines highlight higher metazoan phyla composition (Fig. R6-R7); in fact, during these two campaigns more species are detected and more assigned reads are obtained. The spring and summer campaigns yield lower species richness and reads. The species composition in each season survey is nearly the same for both pipelines.

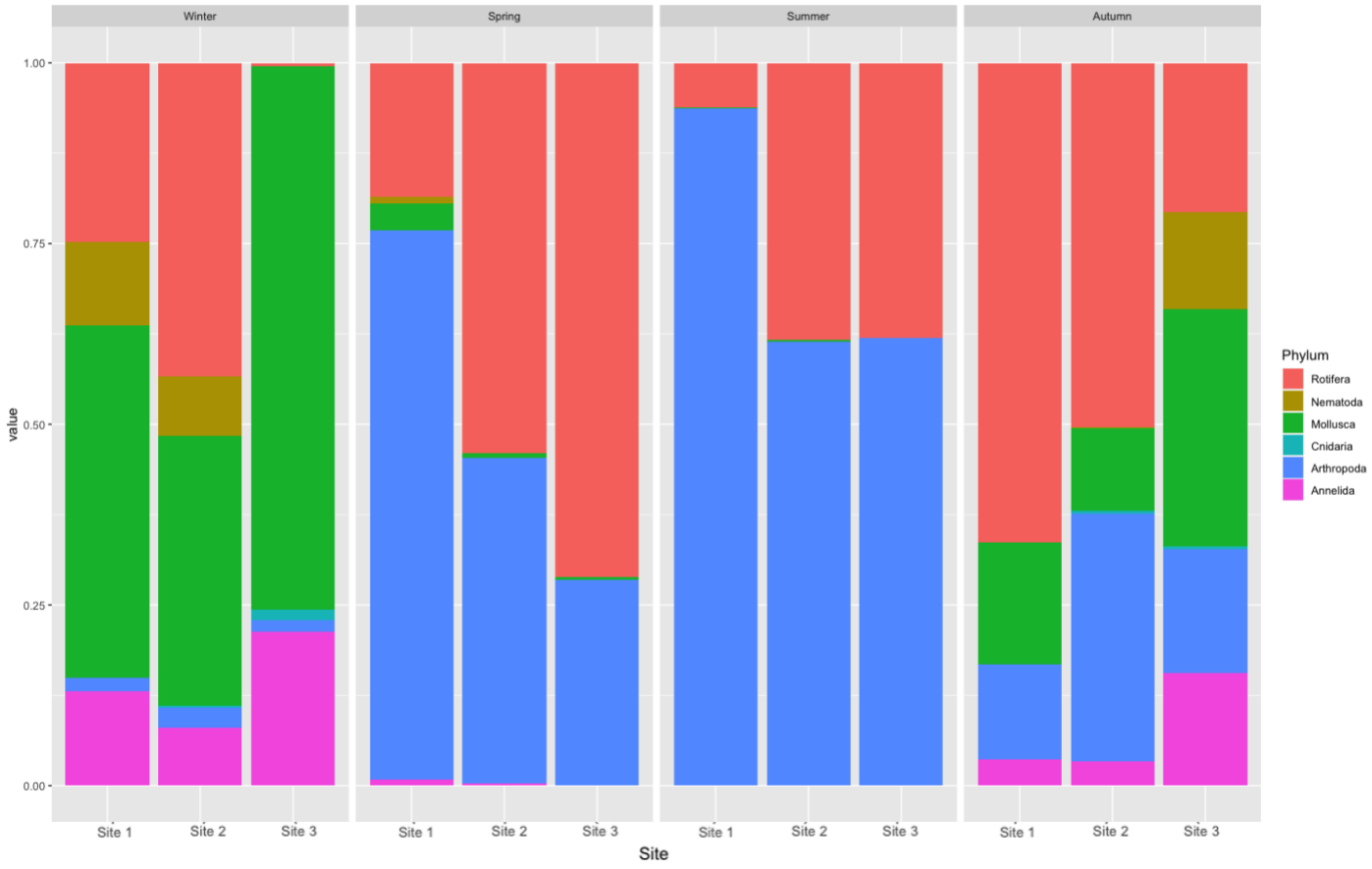


Figure R6. Barplots showing invertebrate phyla composition at each site for each season of 2020 for the FWH marker (OBITools pipeline)

Figure R7. Barplots showing invertebrate phyla composition at each site for each season of 2020 for the FWH marker (QIIME2 pipeline)

3.4 DISCUSSION

In this section, I will consider step by step the results obtained during both the preliminary test and the 2020 campaign; in details I will:

- present how the preliminary 2019 survey has been useful in order to adjust and improve the 2020 campaign
- analyze the fish community results from 2020 campaign comparing the data from the QC with the traditional surveys and the results from the OBITools and the QIIME2 pipelines
- compare the data from the electrofishing method with the metabarcoding results
- discuss the 2020 invertebrate community results from the OBITools and QIIME2
- compare the data from the traditional methods and the metabarcoding results

3.4.1 Preliminary study

The preliminary study was conducted to test different filtering techniques, both in the field and in the laboratory, so as to find the most suitable, efficient and practical protocol to carry on during the season samplings in 2020. This first study also allowed to verify the success of the wet laboratory practices such as the PCRs results and the correct sequencing and chemistry of the library for the MiFish primer. The results not only allowed to adjust and choose the most reliable techniques and lab protocols, they also provided data to be compared to those from the 2020 summer campaign.

During the preliminary study half of the water samples collected from Scolo Dosolo were pre-filtered in the field using a hand pump. This operation was extremely time consuming, since it also implied the relocation of all the equipment and devices from site to site rising the chance of samples contamination.

All the species detected in the pre-filtered samples were included in the non-prefiltered ones. On the contrary, four species (*Abramis brama, Alburnus alburnus, Ameiurus melas* and *Dicentrarchus labrax*) retrieved in the non-prefiltered samples were absent in the pre-filtered ones (Fig. A.)

Another consequence of the pre-filtration step, even if predictable, is the fact that the pre-filtered samples showed a minor number of reads compared to the only once filtered ones. This also might explain the loss of the species described before.

For these reasons, in the 2020 campaigns the prefiltering step was removed in order to prevent possible false negatives and to make more practicable, easy and rapid the collection and analyses of samples.

3.4.2. Fish community 2020 campaign: metabarcoding

From here on, as I stated before, I will mainly consider data from the 2020 campaign for standardization and homogeneity reasons since the results from the 2019 test were obtained with slightly different approaches and methods.

OBITools

All the species from the traditional survey were identified in the OBITools metabarcoding protocol with the exception of two species: *Lepomis gibbosus* and *Alburnus arborella.* (Fig D1)

Lepomis gibbosus is present in the QC, but during the electrofishing sessions only one specimen was found during the summer survey. Its presence could be very limited and therefore the detection using eDNA difficult.

The same could apply to *Alburnus arborella,* in fact only one individual was caught and released during the summer campaign.

During the metabarcoding protocol MOTUs assigned to the Leuciscinae subfamily or to *Leuciscus leuciscus* are retrieved; on the whole, this could probably represent a record from an individual of a slightly genetic different population of *Leuciscus leuciscus.* This species, in fact, is not reported in Italy, and could be therefore a misidentification.

One last point to be discussed regards *Ameiurus* and *Gambusia,* the *ecotag* command assigned one MOTU to *Ameiurus nebulosus* and one to *Gambusia affinis* while in the QC and 2019 survey

Ameiurus melas and *Gambusia holbrooki* were reported. It is almost certain that we are dealing with some sort of misidentification, since *Gambusia affinis* is not reported and *Ameiurus nebulosus,* even if introduced in Italy, is less common than *Ameiurus melas.*

QIIME2

In the QIIME2 species list 4 species are present in the electrofishing survey, in the QC and in the OBITools list: *Ameiurus* sp., *Carassius* sp., *Gambusia* sp. and *Pseudorasbora parva* (Tab. D1).

Here, *Ameiurus* and *Gambusia* are identified only at the genus level but they could well correspond to *Ameiurus melas* and *Gambusia holbrooki* reported in the QC and found during the traditional fish survey.

With the exclusion of the summer sampling, *Carassius gibelio* was also listed in the QIIME2 output and it never appeared during any survey. Since the species belonging to the *Carassius* genus are often subjected to the phenomenon of hybridization (Kottelat & Freyhof, 2007), its presence it is not be excluded in Scolo Dosolo.

Even the other three species listed in the QIIME2 output (but absent in the electrofishing survey and QC) are all reported in Northen Italy. *Esox lucius* was found only in winter and autumn, this possibly being linked to the fact that they tend to move to deeper and cooler water in summer. *Oncorhynchus mykiss* is one of the 100 worst invasive species in the world (Lowe et al., 2000). That being said, its presence in the area has never been reported. Since it is a very common species eaten all around Italy, its detection is more likely to be linked to some sort of contamination from urban waters flowed into the canal.

In Italy there are three different endemic species of *Rutilus,* but the autochthonous *Rutilus rutilus* is well acclimatized and present in the Po basin. It is a ubiquitous and very adaptable species, sometimes, in fact, it could also impact on population number of autochthonous species as it happened for *Alburnus arborella* in Lake Maggiore and Lake Lugano (CISPP 2005). This could also explain the low density of the two autochthonous population of *Alburnus arborella* and *Squalius cephalus* and even the absence of the latter species in the eDNA survey.

The presence of *Ameiurus* sp., *Gambusia* sp. (most likely *Ameiurus melas* and *Gambusia holbrooki), Carassius* sp. and *Pseudorasbora parva* is confirmed in all metabarcoding pipelines both from 2019 and 2020. *Cyprinus carpio* is also present in all the reports but absent in the QIIME2 2020 list. In the 2019 metabarcoding list no MOTU were assigned to this species but a Ciprininae MOTU is given.

Electrofishing vs *eDNA*

All species reported in the QC were also found during the 2020 electrofishing survey with the exception of *Squalius cephalus* (Tab. D1). On the other hand, during the fall 2020 sampling *Ameiurus melas,* which was not reported in the QC document yet, was caught at all sites.

All fishes found during the electrofishing sessions were retrieved in the corresponding eDNA metabarcoding list (Tab. D1). The taxonomic assignment performed by *ecotag* identified *Ameiurus melas* while during the traditional survey the identification was at the genus level. On the other hand, *Pseudorasbora parva* found during the electrofishing was assigned only at the genus level and *Cyprinus carpio* could correspond to the Cyprininae MOTU identified in the eDNA list. In the metabarcoding results five more species were present but not found through the traditional survey.

Table D1. Summary table with all the results for the fish community surveys in comparison. Columns "e" stands for electrofishing, "o" stands for OBITools and "q2" for QIIME2. In blue the species reported in the QC are highlighted.

	Summer 2019						Winter 2020						Spring 2020						Summer 2020									Autumn 2020								
	Site 1		Site 2		Site 3		Site 1		Site 2		Site 3		Site 1		Site 2		Site 3		Site 1			Site 2			Site 3			Site 1			Site 2			Site 3		
Species	**e**	**o**	**e**	**o**	**e**	**o**	**o**	**q2**	**o**	**q2**	**o**	**q2**	**o**	**q2**	**o**	**q2**	**o**	**q2**	**e**	**o**	**q2**	**e**	**o**	**q2**	**e**	**o**	**q2**	**e**	**o**	**q2**	**e**	**o**	**q2**	**e**	**o**	**q2**
Abramis brama		x																																	
Alburnus alburnus		x		x																															
Alburnus alrborella		x				x													1																
Ameiurus melas		x				x																						4			1			2		
Ameiurus nebulosus							x		x		x		x													x			x			x			x	
Ameiurus sp.			1											x				x			x					x			x							
Carassius sp.		x	9	x	14	x	x	x	x	x	x	x	x	x	x	x	x	x	0	x	x	1	x						x	x		x			x	x
Carassius gibelio												x				x																	x			
Cyprinius carpio	22		5				x		x		x		x		x		x		3	x		5	x		2	x		1	x		2	x		4	x	
Cyprininae		x		x		x																														
Esox lucius												x																								x
Dicentrachus labras						x																														
Gambusia affinis							x		x		x		x				x						x			x			x			x			x	
Gambusia holbrooki		x		x		x													3			12			3			6			1					
Gambusia sp.								x		x		x		x		x		x			x			x			x									x
Gobio gobio						x																														
Lepomis gibbosus																									1												
Leuciscinae							x																													
Leuciscus leuciscus							x		x		x		x		x		x			x			x			x			x						x	
Oncorhynchus mykiss								x		x		x				x		x			x						x			x						x
Pseudorasbora parva	4		3		2		x	x	x		x	x	x	x	x	x	x	x	11	x	x	5	x		7	x	x	54	x		11	x		16	x	x
Pseudorasbora sp.		x		x		x																														
Rutilus sp.								x				x		x		x		x			x						x			x						x
Squalius cephalus																																				

Abramis brama is a freshwater fish belonging to the Cyprinidae family and it is an allochthonous species introduced in Italy in the XX century. It is found only in the site 1 filtered sample, so it could be possible that some genetic material has flown into the gutter from other near water courses or upstream. That being said, its presence in the Po valley is well known and confirmed since it has started to compete for feeding resources with *Cyprinus carpio* and *Carassius* spp. drastically reducing their populations both globally and locally (Kottelat & Freyhof, 2007).

Alburnus alburnus and *Alburnus arborella* are congeneric species of the Ciprinidae family. They are morphologically and genetically very similar; for long time *Alburnus arborella* was considered a subspecies of *Alburnus alburnus (*Kottelat & Bianco, 2005). *Alburnus alburnus* is a European species but it is absent in Italy where *Alburnus arborella* is present instead, endemic in the Po valley. The presence in the eDNA list of *Alburnus alburnus* could be a misidentification due to the strict phylogenetic relationship and genetic similarity with *Alburnus arborella.* In fact, the latter

is reported in the Quadro Conoscitivo (QC) - Cassa di espansione Dosolo and found during the 2020 electrofishing campaign but not in the eDNA lists.

Dicentrarchus labrax is a marine and brackish fish, so its presence in the eDNA list is clearly due to some sort of contamination or bioinformatic misidentification. Moreover, it has been detected only in one sample of the preliminary study and never during the 2020 campaign.

Gambusia holbrooki was introduced in Italy and Europe during the XX century as a biological pest control against mosquitos. Its presence is described in the QC and it is also detected during the 2020 traditional and metabarcoding surveys.

Gobio gobio is a freshwater fish autochthonous in the Po basin, it lives in clear shallow waters and it is sedentary. It was present only in the site 3 during the test study but not in the 2020 campaign. Also in this case, some of its genetic material could have been transported from upstream and detected in the Scolo Dosolo.

3.4.3 Freshwater invertebrates 2020 campaign: metabarcoding

OBITools

All the species detected in Scolo Dosolo are reported in Italy or at least in Europe. Only for 4 species there are not records available in Italy or Europe. These are: *Amiphichaeta raptisae,* a North American Annelida found only in the spring campaign; *Caenorhabditis brenneri,*a Nematoda found in the tropics; *Hypoponera opacior,* an American ant indoor introduced only in Norway; *Sulcospira paludiformis* a gastropod known from Hainan (China).

On the basis of present literature data, different hypotheses can be taken into account; it is not possible in fact to define if we are dealing with false positives, wrong taxonomic assignation or if these are the first record of these animals' effective presence in our territory. Obviously, further investigations are needed to verify the presence of these species in Scolo Dosolo using group-specific sampling strategies.

In the OBITools list, 22 species absent in the QIIME2 output were found. On the whole, this pipeline retrieves 15 species more than QIIME2 approach (Tab. D2).

Table D2. Summary table with all the results for the invertebrate community surveys in comparison. Column "o" stands for OBITools and "q2" for QIIME2.

	Winter 2020						Spring 2020						Summer 2020						Autumn 2020					
	Site 1		Site 2		Site 3		Site 1		Site 2		Site 3		Site 1		Site 2		Site 3		Site 1		Site 2		Site 3	
Species	**o**	**q2**	**o**	**q2**	**o**	**q2**	**o**	**q2**	**o**	**q2**	**o**	**q2**	**o**	**q2**	**o**	**q2**	**o**	**q2**	**o**	**q2**	**o**	**q2**	**o**	**q2**
Acanthocyclops americanus	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Aculops lycopersici	x				x						x												x	x
Amphichaeta raptisae			x																					
Anuraeopsis sp.	x	x	x	x	x						x		x	x	x	x	x	x	x	x	x	x	x	x
Asplanchna sieboldi	x	x	x		x		x				x								x	x	x	x		
Balclutha frontalis																								
Brachionus angularis	x						x		x			x							x		x	x	x	x
Brachionus quadridentatus			x	x	x	x	x		x		x	x	x	x	x	x	x	x	x	x	x	x	x	x
Brachionus sp.										x										x				
Brachionus variabilis							x		x		x	x							x	x			x	x
Branchiura sowerbyi			x	x	x	x													x		x	x	x	x
Caenorhabditis brenneri	x	x	x	x			x				x	x									x		x	x
Chaetogaster diastrophus	x	x	x	x	x	x															x		x	x
Chironomus sp.					x		x		x		x										x		x	
Chironomus plumosus												x												
Chironomus transvaalensis			x	x	x	x	x				x								x	x	x	x	x	
Cricotopus sp.							x		x	x	x													
Culicidae																			x					
Daphnia magna							x	x			x	x												
Dero digitata			x		x														x	x	x	x	x	x
Dero sp.						x														x				
Deroceras reticulatum		x																		x				
Deroceras sp.	x		x										x						x					
Diplogasteridae		x		x		x												x		x				
Dicrotendipes nervosus																							x	
Eucyclops cf. *estherae*					x																			
Eucyclops sp.						x																		
Formica exsecta	x		x																x	x	x	x	x	x
Habrotrocha sp.		x		x																				
Globulidrilus riparius																							x	
Glyptotendipes sp.	x																							
Hydra vulgaris			x	x	x	x											x		x		x	x	x	x
Hypogastrura vernalis																			x	x	x	x		
Hypoponera opacior																			x					
Keratella cochlearis	x		x	x			x	x	x	x	x	x	x				x	x	x	x				
Keratella tropica				x																		x		
Keratella sp.			x																x		x			
Limnodrilus hoffmeisteri	x		x		x	x													x	x	x	x	x	
Limnodrilus sp.	x	x	x	x	x	x	x		x										x	x	x		x	
Liposcelis rufa			x		x														x		x	x		
Macrothrix sp.	x						x	x	x	x	x	x	x	x	x		x	x						
Myrmica sp.					x														x		x		x	
Myrmica sabuleti																						x		x
Nais barbata	x	x	x	x	x	x																		
Nais christinae	x		x																					
Nais communis/variabilis complex sp.	x	x	x	x	x	x	x		x		x		x								x	x	x	
Nais variabilis		x		x																				
Nylanderia sp.	x						x		x								x							
Othius laeviusculus											x						x		x	x				
Paranais frici	x	x	x	x	x	x													x				x	
Phragmatobia fuliginosa																							x	
Physella acuta	x	x	x	x	x	x	x	x	x		x	x	x	x	x		x		x	x	x	x	x	x
Polyarthra dolichoptera	x	x	x	x	x	x	x	x			x	x	x	x	x	x	x	x	x	x	x	x	x	
Polyarthra dolichoptera complex sp.	x		x		x		x		x		x		x						x		x		x	
Polyarthra sp.															x									
Polyarthra vulgaris				x														x						
Procambarus clarkii																			x					
Potamothrix bavaricus						x														x		x		x
Psammoryctides barbatus	x	x			x														x	x	x			
Scatopsciara vitripennis			x																					
Rhabditida								x																
Solenopsis sp.																				x		x		
Solenopsis fugax																			x		x		x	
Stylodrilus lemani	x																							
Sulcospira paludiformis	x	x	x	x	x	x											x		x		x	x	x	x
Tabanus autumnalis																			x					
Thermocyclops cf. *taihokuensis*	x																				x		x	
Tubifex tubifex	x	x	x		x	x															x		x	
Xiphinema brevicollum	x																							

The lower species composition in metazoan taxa during the spring and summer surveys, noted in the results section, could be due to the fact that the water was full of algae and debris. If we consider also the reads assigned to non-metazoan groups such as Oomycota, Discosea, Ascomycota, and Loukozoa we notice that the majority of the sequences belongs to these taxonomic groups (Fig. D1). The presence of algae drove a preferential PCR amplification of these taxonomic group because the primers used have also high affinity for algae and other eukaryotic taxa. (Leese et al., 2021). During the cold seasons, these groups are less present in the water (Lanciotti et al., 2003), possibly allowing the detection of more animal taxa.

Figure D1. Barplots showing eukaryotic phyla composition at each site for each season of 2020 for the FWH marker (OBITools pipeline)

The fresh water eDNA has been able to detect only 14 terrestrial species among the 54 listed (Tab. D2). They belong all to the Arthropoda phylum, with the exception of the gastropod *Deroceras* sp. The presence in water of genomic material belonging to these species could have been traced back to secondary predation (Hardy et al., 2017) or also to dead individuals fallen into the canal.

QIIME2

All the species reported in the QIIME2 list are reported in Italy or at least in Europe. Even in this pipeline two AVSs are assigned to *Caenorhabditis brenneri* and *Sulcospira paludiformis* whose reported distribution at our knowledge does not comprise Italy or Europe.

A total of 11 species are present in this list but not in the OBITools output, all likely to be present in Italy and in an environment such as the Scolo Dosolo one.

The same observation stated for the taxonomic association bias in the warmer season is also valid with this different pipeline as can be seen in Figure D2. The two graphics (Fig D1 and D2), in fact, are nearly the same.

The same terrestrial taxa found in the OBITools list are also recovered in the QIIME2 pipeline.

To be noted that 5 MOTUs assigned at the genus level by the OBITools pipeline are assigned at the species level by the QIIME2 pipeline (i.e. *Chironomus plumosus, Deroceras reticulatum, Keratella tropica, Myrmica sabuleti, Polyarthra vulgaris).*

On the other hand, only one ASV identified at genus level is identified at species level by the OBITools piepeline, that is *Solenopsis fugax*.

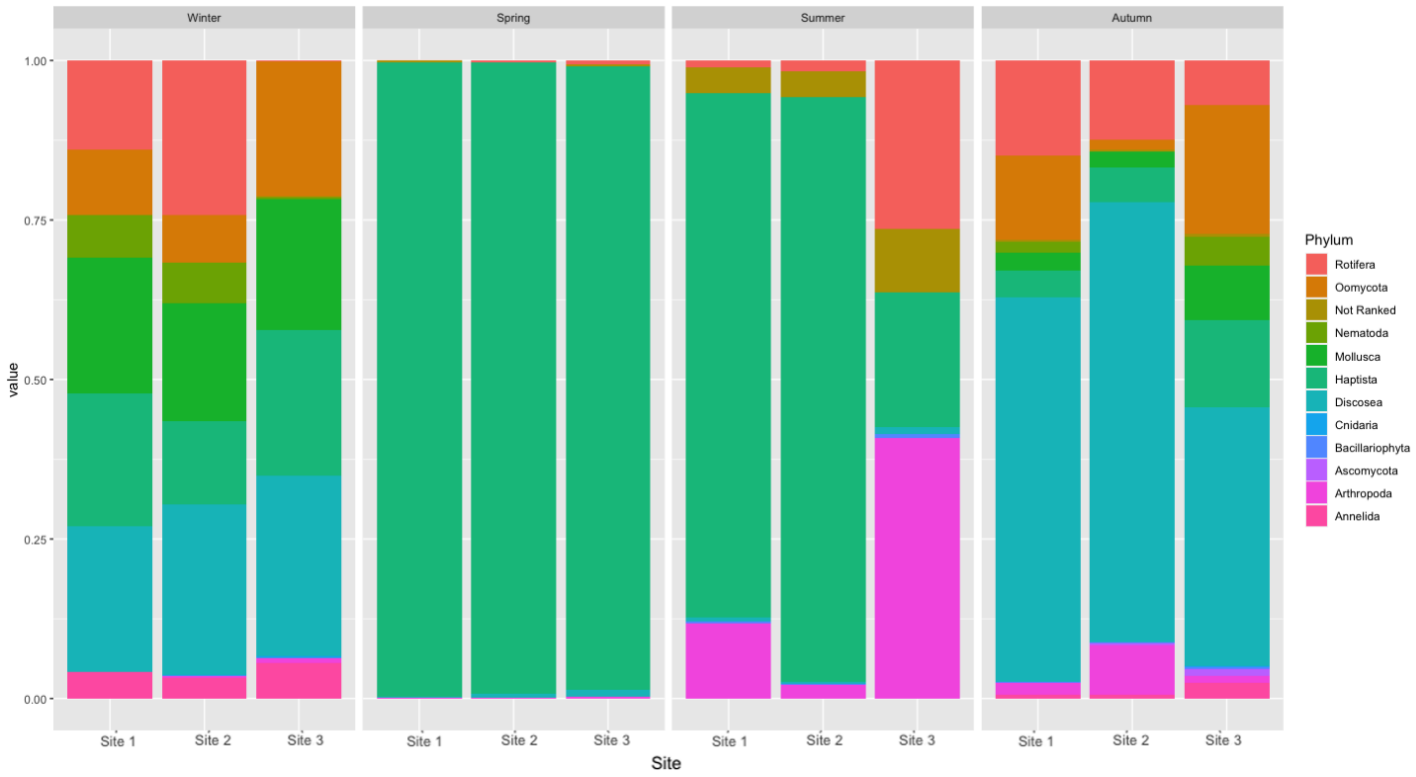


Figure D2. Barplots showing eukaryotic phyla composition at each site for each season of 2020 for the FWH marker (QIIME2 pipeline)

Invertebrates: Traditional vs *metabarcoding approach*

All the morphological identifications of the invertebrate samples were not performed by specialized researcher, but the molecular identification was at least always correct at the family or order level. In two cases the identification was incorrect at the species level and only in other two cases they were correct at species level.

None of the species collected with the net and surber sampler were found in the environmental DNA list, with the exception of *Procambarus clarkii* found in the output of the OBITools pipeline.

This discrepancy between the results from traditional methods and metabarcoding could be due to the fact that all the insects collected using the entomological net were strictly terrestrials with the exception of the Odonata species which have aquatic nymphs. Moreover, the species caught

with the surber sampler were benthic nymphs or aquatic insects and it is a known fact that insects and arthropods in general release little quantity of genetic material in water (see Introduction). Further, Odonata nymphs tend to live near the bottom of the water courses, so it is unlikely to detect their DNA in the surface water in semi-lotic systems such as man-made canals where the phenomenon of resuspension is not predominant.

The majority of the taxa found thanks to the metabarcoding technique is microscopic, therefore they can be detected and identified only using *ad hoc* protocols analyzing water samples under the microscope. Moreover, a very specialized knowledge is needed to identify such taxa at species level. This is also the main reason why there are little or no information about the presence and distribution of these taxa in any official regional or national freshwater surveys. In fact, in the QC *Procambarus clarkii* is the only invertebrate reported probably because it is an invasive alien species whose presence is very easy to determine. The eDNA metabarcoding can be source of a lot of new information about overlooked taxa that can also provide insight about the quality of water and habitats.

3.4.4 eDNA metabarcoding protocol and cost

Defining a standardized protocol in eDNA metabarcoding survey is complicated, since each protocol needs to be adjusted to a series to factors and parameters typical to a certain environment and aim of the project itself.

However, after the results presented here, it Is possible to suggest a protocol suitable for internal freshwater courses in the Po Basin.

First of all, the sampling sites should be identified based on the length and hydrogeological characteristics of the water course, but they should not be less than three. A complete seasonal sampling it also recommended to possibly detect the presence of animals along the year.

A prefiltering step on field is discouraged since, during this study, it has been the source of many drawbacks. In fact, this practice led to a lesser number of reads and species detected, not to

mention the fact that it implied the relocation of the equipment for this operation at all sampling sites making the survey more time consuming.

It is strongly recommended to collect water samples using sterilized 1000 mL Nalgene bottles and perform a single filtering step in laboratory using enclosed filter capsule in order to limit the possibility of contamination.

The extraction protocol proposed by Spens et al. 2017 using the DNeasy Blood & Tissue kit (Qiagen) is suggested since it is specifically designed for freshwater environmental samples, it is pretty straightforward to follow and it has produced good results.

For the amplification step, the AmpliTaq Gold 360 Master Mix (Applied Biosystems) allows a safer handling of the samples being a hot start polymerase.

Obviously, the primer choice is a crucial step in every metabarcoding survey. With this study, the MiFish set of primers is confirmed to hold good results also on freshwater fishes. On the other hand, when dealing with broader taxa like arthropods o invertebrates it is strongly suggested to perform in silico analyses using different set and primer combination.

For the library preparation, the Metafast protocol by Fasteris SA is designed to manage amplicons from eDNA PCR limiting tag jumps and therefore it is worth to take it in consideration if a single-step PCR with tagged primer is used for multiplexing samples (see chapter 2.6.2 Multiplexing samples).

The paired-end sequencing on Illumina MiSeq is one of the most used sequencing technologies when dealing with eDNA metabarcoding samples. Said that, thanks to the constant innovation in this sector, the NovaSeq 6000 may be a better instrument for environmental metabarcoding studies (Singer et al., 2019).

Regarding the bioinformatic pipeline, both ObiTools and QIIME2 have their pro and cons. If one has to choose between the two, QIIME2 has turned out to be faster and more user-friendly to use. It was able to retrieve more fish species, and even though in the case of *Ameiurus* and

Gambusia the species level was not defined, at least it did not lead to a species misidentification in contrast with ObiTools (see chapter 3.4.2. Fish community 2020 campaign: metabarcoding). For invertebrates even if in this case less species were retrieved contrary with ObiTools, more ASVs were assigned at the species level in comparison to ObiTools. Moreover, only two species present in the ASVs table were not reported in Europe, whereas in the ObiTools list four non-European taxa were reported.

The most expensive step was definitely the library preparation and the amplicons sequencing, Fasteris SA demanded about €325 for the preparation of a single amplicon library and about €1000 for the sequencing. All the equipment for the filtration and reagent for the extraction and amplification amounted to €50 ca. per samples. To sum up, a single sample, which in this case corresponds to a sampling site, with 20 replicates (depending on how one will arrange the PCR plate and how many controls will be included), will cost €160. The single sample cost could be reduced to €80 arranging the PCR plate so that the number of biological samples are increased and the replicates are lowered.

3.4.5 Risk assessment and decision making

A rapid and effective tool such as the eDNA metabarcoding is crucial in the early detection of alien invasive species. Having outlined and validated a protocol suitable for the internal water courses could be helpful in collecting data on such species in order to develop, in future, risk maps of the possible distribution of species of interest. The eDNA technology permits to work on a broad spatial range, allowing a quick detection and collection of data on possible distribution of invasive species or species linked to risk-management. Many invasive species, like *Dreissena polymorpha*, represent a threat to hydraulic security so that the knowledge of the potential distribution can be fundamental in shaping intervention strategies.

Creating a risk map thanks to the collection of data deriving from eDNA surveys will represent an effective tool to focus on the sites most likely to be threaten by invasive species, species vectors of pathogens or species dangerous for the water course security.

3.5 FINAL REMARKS

From its first application until today the metabarcoding of environmental DNA is still showing all its power and potentiality alongside with its limits.

First of all, at the present time limited biodiversity investigation on freshwater systems have been conducted in Italy. A recent work by Jamwal et al. 2021 assessing the presence of the European otter in southern Italy using a qPCR on eDNA has been published. Other research targeted fungi in bioaerosol across Northern and Central Italy (Tordoni et al., 2021), plant diversity in alpine glacier cores (Varotto et al., 2021) and marine biodiversity in the vicinity of the three offshore gas platforms in North Adriatic Sea (Cordier et al., 2019). This is actually the first study, using a multi-taxa approach, to assess the biodiversity in a man-made canal in Italy.

Many researches were focused on the optimization of sampling strategies of running water (Balasingham et al., 2017) and still water bodies (Hänfling et al., 2016), whereas only few studies investigated how water dynamics could influence the detection of species in semi-lotic and regulated flow environment such as man-made canals (McDevitt et al., 2019). Here it is shown that the water movements in canals are enough to reduce the number of samples and operate in an environment similar to a lentic system since with only three sampling sites along the canals all the species known to be in Scolo Dosolo from the QC are present in metabarcoding list, with the exception of *Lepomis gibbosus* (one specimens found during the summer 2020 electrofishing session) and *Squalius cephalus* (which could be possible no longer present in the canal)

The decision of carrying out a preliminary study has been fundamental in order to have a first knowledge of the characteristics of the habitat and to adjust and improve laboratory steps. A pilot test is deeply needed before investing time and resources on the main research on eDNA. In this way it is possible to try out different sampling strategies, different eDNA concentration and isolation method, appropriate primer choice and success of PCR, right sequencing chemistry and sequencing depth. In this case the preliminary study conducted during the summer 2019 allowed to discard the prefiltering step leading to the detection of more animal species and

probable to limit the potential contamination in the field. Moreover, the data obtained were used to verify species presence between the summer 2019 campaign and the summer 2020 campaign.

The metabarcoding of eDNA provides different pros if compared to traditional and conventional methods.

First of all, it is a time saver approach since one has only to collect the environmental samples with all the anti-contaminations precautions. On the other hand, all the traditional sampling methods take lots of time to be performed and carried on. In this study, the water sampling took only from 15 to 20 minutes to be concluded in each site while the electrofishing took up to 40-50 minutes for site, which is almost three time longer.

As a consequence, the method results more practical and handier. Once the equipment is treated to avoid all the possible contamination, only two bottles of water per site need to be collected. On the contrary, the electrofishing needed a long time of preparation of the equipment, the search of a suitable point which allowed the placement in water of the boat and the packing of the gears at the end.

Given that it is necessary to collect only the environment samples in the field, everyone can do it, even people without any technical or specific skills. This fact led to the realization of citizen science-based monitoring projects mostly focused on biodiversity assessment (Biggs et al., 2015; Larson et al., 2020). Most of the traditional methods, instead, imply technical knowledge and skills limited to scholars or researchers.

The eDNA methods, further, are non-invasive approaches. The presence of species is not linked to their collection and manipulation. Moreover, their habitats are not disrupted and they are almost untouched. A direct consequence is that, most of the time, no permissions are needed for handling and dealing with organisms which can be also protected by law. On the other hand, traditional methods as the electrofishing technique or the entomological net and surber sampler need the pick-up and/or the collection of organisms. In fact, for this study all the authorization

were made for handling of fish species. An inconvenience to this mandatory step resulted in the delay of all the sampling for this study.

Another advantage linked to the eDNA methods is the reduction of the possibility to spread alien/invasive/allochthonous species since there is no handling of organisms. The spread of pathogens is also lowered because all the gears and equipment used during the samplings need to be treated and sanitized to avoid contaminations.

The identification based on morphology can be very difficult at the species level. This is even harder on juvenile forms, larvae, nymphs and eggs since most of the taxonomic keys focus on adult specimens. Further, and unfortunately, taxonomic knowledge is becoming more and more scarce and taxa-specialists are slowing lowering in number. Once all the precautions are taken, the metabarcoding of eDNA can provide high taxonomic resolution.

It is to be noted that, in order to check the presence of false positives but especially false negatives, a good knowledge of the biodiversity of the area of interest should be available. For Emilia-Romagna region a lot of information is already available: this comes in particular from traditional campaigns performed in Natura 2000 protected areas. For Scolo Dosolo a QC was also available and it came very handy for species presence comparison. A powerful tool such as the eDNA could be taken therefore into consideration to support traditional data and could be helpful when dealing with vast areas. eDNA projects could well be developed in our region owing to the richness in data coming from all the traditional reports conducted in protected areas and regional parks.

Obviously, all methods come with drawbacks and pitfalls when compared with traditional and conventional technique.

The first con is that eDNA method is a qualitative method which basically gives information only about the presence or the absence of a given taxon. The eDNA metabarcoding technique does not reveal anything about life stages, size, age class distribution, health or fecundity of the target

organisms. This could be crucial, for example, when dealing with invasive alien species or threatened organisms.

Even though more and more studies are focusing on finding a direct relationship between the density of a species and the quantity of genetic material released into the environment (Takahara et al., 2012; Lacoursière-Roussel et al., 2016), under field conditions many factors influence the presence and persistence of eDNA (Barnes et al., 2014). Only when using a qPCR method on single species, information about the abundance of the target species is available, otherwise eDNA can give only some indications about trends in densities at the moment. Nevertheless, it should be considered that also traditional methods are often not well implemented for good densities estimates either (Taberlet et al., 2018).

Since mitochondrial DNA abundance in the environment is higher if compared to nuclear DNA, most of the eDNA protocols use mitochondrial markers. In this way, it is extremely difficult to distinguish between maternal species and hybrids (Fukumoto *et al.,* 2015). Traditional methods based on morphology can sometimes succeed in the identification of hybrids, even though it can be difficult also in this case if hybrids look like one of their parental species.

Surely, one of the main concerns regarding eDNA methods is the management of contaminations. All the drawbacks that characterize PCR protocols are present in eDNA protocols. One should deal with cross-contamination, PCR stochasticity, failure in DNA extraction or amplification. For this reason, all the anti-contamination precautions should be taken and followed from start to finish.

One final point concern specifically this work. The majority of surveys took place during the 2020, this meant working following the safety guidelines issued by the Italian government to prevent the spreading of the Covid-2019 virus. This situation did not allow to carry out the traditional faunistic surveys during spring 2020, while the water samplings were conducted anyway. Since the eDNA metabarcoding protocols relies on the analysis of environmental samples, as bottles of water in this case, the study kept going even in these limiting circumstances.

To conclude, the eDNA metabarcoding technology is becoming each day more studied and investigated since it has already been revealed as a powerful tool. Thanks to the contributions of many and many researchers sampling protocols, extractions methods, primer availability, sequencing technologies and bioinformatic pipelines are improving and being refined. Seeing all the pros and cons presented above, the fundamental result of this work regards the desirability of an integration between eDNA metabarcoding, traditional methods and ecological a taxonomical expertise. Given that the aim is the achievement of a faster, reliable and efficient monitoring and assessment for species and ecosystems, one should see the eDNA not as a replacement to traditional monitoring methods but as further resource to implement our knowledge of biodiversity, fundamental for its conservation.

4 BIBLIOGRAPHY

Allen, M. C., Nielsen, A. L., Peterson, D. L., & Lockwood, J. L. (2021). Terrestrial eDNA survey outperforms conventional approach for detecting an invasive pest insect within an agricultural ecosystem. *Environmental DNA*.

Andersen, K., Bird, K. L., Rasmussen, M., Haile, J., Breuning-Madsen, H. E. N. R. I. K., Kjaer, K. H., ... & Willerslev, E. (2012). Meta-barcoding of 'dirt'DNA from soil reflects vertebrate biodiversity. *Molecular Ecology*, 21(8), 1966-1979.

Andrews, S. (2010). FastQC: A Quality Control Tool for High Throughput Sequence Data [Online]. Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/

APAT-IRSA/CNR (2004). Manuali e linee guida 29/2003. Metodi analitici per le acque, vol.3; 9010 Indice Biotico esteso (I.B.E.)

Apothéloz-Perret-Gentil, L., Cordonier, A., Straub, F., Iseli, J., Esling, P., & Pawlowski, J. (2017). Taxonomy-free molecular diatom index for high-throughput eDNA biomonitoring. *Molecular ecology resources*, 17(6), 1231-1242.

Balasingham, K. D., Walter, R. P., & Heath, D. D. (2017). Residual eDNA detection sensitivity assessed by quantitative real-time PCR in a river ecosystem. *Molecular Ecology Resources*, 17(3), 523-532.

Barnes, M. A., & Turner, C. R. (2016). The ecology of environmental DNA and implications for conservation genetics. *Conservation genetics*, 17(1), 1-17.

Barnes, M. A., Turner, C. R., Jerde, C. L., Renshaw, M. A., Chadderton, W. L., & Lodge, D. M. (2014). Environmental conditions influence eDNA persistence in aquatic systems. *Environmental science & technology*, 48(3), 1819-1827.

Bayer SR, Countway PD, Wahle RA (2019) Developing an eDNA toolkit to quantify broadcast spawning events of the sea scallop Placopecten magellanicus: moving beyond fertilization assays. *Mar Ecol Prog Ser* 621:127-141.

Bazinet, A. L., & Cummings, M. P. (2012). A comparative evaluation of sequence classification programs. *BMC Bioinformatics*, 13(1), 1-13.

Bellemain, E., Davey, M. L., Kauserud, H., Epp, L. S., Boessenkool, S., Coissac, E., ... & Brochmann, C. (2013). Fungal palaeodiversity revealed using high-throughput metabarcoding of ancient DNA from arctic permafrost. *Environmental microbiology*, 15(4), 1176-1189.

Beng, K. C., & Corlett, R. T. (2020). Applications of environmental DNA (eDNA) in ecology and conservation: opportunities, challenges and prospects. *Biodiversity and Conservation*, 29(7), 2089-2121.

Biggs, J., Ewald, N., Valentini, A., Gaboriaud, C., Dejean, T., Griffiths, R. A., ... & Dunn, F. (2015). Using eDNA to develop a national citizen science-based monitoring programme for the great crested newt (*Triturus cristatus*). *Biological Conservation*, 183, 19-28.

Bolyen, E., Rideout, J. R., Dillon, M. R., Bokulich, N. A., Abnet, C. C., Al-Ghalith, G. A., ... & Caporaso, J. G. (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nature biotechnology*, 37(8), 852-857.

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.

Buermans, H. P. J., & Den Dunnen, J. T. (2014). Next generation sequencing technology: advances and applications. Biochimica et Biophysica Acta (BBA)-*Molecular Basis of Disease*, 1842(10), 1932-1941.

Bylemans, J., Gleeson, D. M., Hardy, C. M., & Furlan, E. (2018). Toward an ecoregion scale evaluation of eDNA metabarcoding primers: A case study for the freshwater fish biodiversity of the Murray–Darling Basin (Australia). *Ecology and evolution*, 8(17), 8697-8712.

Cavaliere, M., Angeles, I. B., Montresor, M., Bucci, C., Brocani, L., Balassi, E., ... & Frontalini, F. (2021). Assessing the ecological quality status of the highly polluted Bagnoli area (Tyrrhenian Sea, Italy) using foraminiferal eDNA metabarcoding. *Science of The Total Environment*, 147871.

CISPP. (2005). Ricerche ed attività della Commissione italo - svizzera per la pesca nel quinquennio 2000-2004

Civade, R., Dejean, T., Valentini, A., Roset, N., Raymond, J. C., Bonin, A., ... & Pont, D. (2016). Spatial representativeness of environmental DNA metabarcoding signal for fish biodiversity assessment in a natural freshwater system. *PloS one*, 11(6), e0157366.

Clusa, L., Ardura, A., Gower, F., Miralles, L., Tsartsianidou, V., Zaiko, A., & Garcia-Vazquez, E. (2016). An easy phylogenetically informative method to trace the globally invasive *Potamopyrgus* mud snail from river's eDNA. *PLoS One*, 11(10), e0162899.

Coissac, E., Riaz, T., & Puillandre, N. (2012). Bioinformatic challenges for DNA metabarcoding of plants and animals. *Molecular ecology*, 21(8), 1834-1847.

Cordier, T., Frontalini, F., Cermakova, K., Apothéloz-Perret-Gentil, L., Treglia, M., Scantamburlo, E., ... & Pawlowski, J. (2019). Multi-marker eDNA metabarcoding survey to assess the environmental impact of three offshore gas platforms in the North Adriatic Sea (Italy). *Marine environmental research*, 146, 24-34.

D'Antoni S., Tancioni L., Scardi M., Marino G., (2007). Protocollo di campionamento e analisi della fauna ittica dei sistemi lotici - Manuali e Linee Guida APAT – Metodi Biologici per le acque – Parte I – XX/2007.

Deagle, B. E., Jarman, S. N., Coissac, E., Pompanon, F., & Taberlet, P. (2014). DNA metabarcoding and the cytochrome c oxidase subunit I marker: not a perfect match. *Biology Letters*, 10(9), 20140562.

Deere, D., Porter, J., Pickup, R. W., & Edwards, C. (1996). Survival of cells and DNA of Aeromonas salmonicida released into aquatic microcosms. *Journal of applied bacteriology*, 81(3), 309-318.

Deiner, K., & Altermatt, F. (2014). Transport distance of invertebrate environmental DNA in a natural river. *PloS one*, 9(2), e88786.

Deiner, K., Bik, H. M., Mächler, E., Seymour, M., Lacoursière-Roussel, A., Altermatt, F., ... & Bernatchez, L. (2017). Environmental DNA metabarcoding: Transforming how we survey animal and plant communities. *Molecular Ecology*, 26(21), 5872-5895.

Dejean, T., Valentini, A., Duparc, A., Pellier-Cuit, S., Pompanon, F., Taberlet, P., & Miaud, C. (2011). Persistence of environmental DNA in freshwater ecosystems. *PloS one*, 6(8), e23398.

Ghetti, P.F. (1997) - Indice Biotico Esteso (I.B.E.). I macroinvertebrati nel controllo della qualità degli ambienti di acque corrente. Provincia Autonoma di Trento. pp. 222.

Edgar, R. C. (2010). Search and clustering orders of magnitude faster than BLAST. Bioinformatics, 26(19), 2460-2461.

Edgar, R. C., Haas, B. J., Clemente, J. C., Quince, C., & Knight, R. (2011). UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics*, 27(16), 2194-2200.

Eiler, A., Löfgren, A., Hjerne, O., Nordén, S., & Saetre, P. (2018). Environmental DNA (eDNA) detects the pool frog (Pelophylax lessonae) at times when traditional monitoring methods are insensitive. *Scientific Reports*, 8(1), 1-9.

Elbrecht, V., Braukmann, T. W., Ivanova, N. V., Prosser, S. W., Hajibabaei, M., Wright, M., ... & Steinke, D. (2019). Validation of COI metabarcoding primers for terrestrial arthropods. *PeerJ*, 7, e7745.

Elbrecht, Vasco, and Florian Leese. (2017).Validation and development of COI metabarcoding primers for freshwater macroinvertebrate bioassessment. *Frontiers in Environmental Science* 5: 11.

Epp, L. S., Boessenkool, S., Bellemain, E. P., Haile, J., Esposito, A., Riaz, T., ... & Brochmann, C. (2012). New environmental metabarcodes for analysing soil DNA: potential for studying past and present ecosystems. *Molecular Ecology*, 21(8), 1821-1833.

Evans, N. T., Olds, B. P., Renshaw, M. A., Turner, C. R., Li, Y., Jerde, C. L., ... & Lodge, D. M. (2016). Quantification of mesocosm fish and amphibian species diversity via environmental DNA metabarcoding. *Molecular Ecology Resources*, 16(1), 29-41.

Ficetola, G. F., Miaud, C., Pompanon, F., & Taberlet, P. (2008). Species detection using environmental DNA from water samples. *Biology Letters*, 4(4), 423-425.

Fonseca, V. G., & Lallias, D. (2016). Metabarcoding marine sediments: preparation of amplicon libraries. In Marine Genomics (pp. 183-196). *Humana Press*, New York, NY.

Foppen, J. W., Orup, C., Adell, R., Poulalion, V., & Uhlenbrook, S. (2011). Using multiple artificial DNA tracers in hydrology. *Hydrological Processes*, 25(19), 3101-3106.

Fukumoto, S., Ushimaru, A., & Minamoto, T. (2015). A basin-scale application of environmental DNA assessment for rare endemic species and closely related exotic species in rivers: A case study of giant salamanders in Japan. *Journal of Applied Ecology*, 52(2), 358-365.

Geller, J., Meyer, C., Parker, M., & Hawk, H. (2013). Redesign of PCR primers for mitochondrial cytochrome c oxidase subunit I for marine invertebrates and application in all-taxa biotic surveys. *Molecular Ecology Resources*, 13(5), 851-861.

Haas, B. J., Gevers, D., Earl, A. M., Feldgarden, M., Ward, D. V., Giannoukos, G., ... & Human Microbiome Consortium. (2011). Chimeric 16S rRNA sequence formation and detection in Sanger and 454-pyrosequenced PCR amplicons. *Genome Research*, 21(3), 494-504.

Hänfling, B., Lawson Handley, L., Read, D. S., Hahn, C., Li, J., Nichols, P., ... & Winfield, I. J. (2016). Environmental DNA metabarcoding of lake fish communities reflects long-term data from established survey methods. *Molecular Ecology*, 25(13), 3101-3119.

Hao, X., Jiang, R., & Chen, T. (2011). Clustering 16S rRNA for OTU prediction: a method of unsupervised Bayesian clustering. *Bioinformatics*, 27(5), 611-618.

Hartvig, I., Kosawang, C., Kjaer, E. D., & Nielsen, L. R. (2021). Detecting rare terrestrial orchids and associated plant communities from soil samples with eDNA methods. *Biodiversity and Conservation*, 1-23.

He, X., Chen, H., Shi, W., Cui, Y., & Zhang, X. X. (2015). Persistence of mitochondrial DNA markers as fecal indicators in water environments. *Science of the Total Environment*, 533, 383-390.

Hebsgaard, M. B., Phillips, M. J., & Willerslev, E. (2005). Geologically ancient DNA: fact or artefact?. *Trends in Microbiology*, 13(5), 212-220.

Hobbs, J., Round, J. M., Allison, M. J., & Helbing, C. C. (2019). Expansion of the known distribution of the coastal tailed frog, *Ascaphus truei*, in British Columbia, Canada, using robust eDNA detection methods. *PloS one*, 14(3), e0213849.

Höss, M., Jaruga, P., Zastawny, T. H., Dizdaroglu, M., & Paabo, S. (1996). DNA damage and DNA sequence retrieval from ancient tissues. *Nucleic Acids Research*, 24(7), 1304-1307.

Jamwal, P. S., Bruno, A., Galimberti, A., Magnani, D., Krupa, H., Casiraghi, M., & Loy, A. (2021). First assessment of eDNA-based detection approach to monitor the presence of Eurasian otter in southern Italy. Hystrix, *The Italian Journal of Mammalogy*.

Jane, S. F., Wilcox, T. M., McKelvey, K. S., Young, M. K., Schwartz, M. K., Lowe, W. H., ... & Whiteley, A. R. (2015). Distance, flow and PCR inhibition: e DNA dynamics in two headwater streams. *Molecular Ecology Resources*, 15(1), 216-227.

Jerde, C. L., Mahon, A. R., Chadderton, W. L., & Lodge, D. M. (2011). “Sight-unseen” detection of rare aquatic species using environmental DNA. *Conservation letters*, 4(2), 150-157.

Ji, Y., Ashton, L., Pedley, S. M., Edwards, D. P., Tang, Y., Nakamura, A., ... & Yu, D. W. (2013). Reliable, verifiable and efficient monitoring of biodiversity via metabarcoding. *Ecology Letters*, 16(10), 1245-1257.

Kamimura, S., Kozuki, Y., Otani, S., Hirakawa, R., Iwami, K., Takeyama, K., & Yamanaka, R. (2018). Fish diversity detection at port and urban canal area using environmental DNA mettabarcoding. *J. Jap. Soc. Civ. Eng.(Ocean Eng.)*, 74(2), 474-479.

Kolman, C. J., & Tuross, N. (2000). Ancient DNA analysis of human populations. *American Journal of Physical Anthropology: The Official Publication of the American Association of Physical Anthropologists*, 111(1), 5-23.

Kottelat M., Freyhof J., Handbook of European Freshwater Fishes, Cornol (CH), Publications Kottelat, 2007, ISBN 88-7021-299-8.

Kottelat, M., & Bianco, P. G. (2005). On the valid name of the alborella, *Alburnus arborella* (Teleostei: Cyprinidae). Ichthyological Exploration of Freshwaters, 16(2), 179.

Kozich, J. J., Westcott, S. L., Baxter, N. T., Highlander, S. K., & Schloss, P. D. (2013). Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. *Applied and Environmental Microbiology*, 79(17), 5112-5120.

Lacoursière-Roussel, A., Côté, G., Leclerc, V., & Bernatchez, L. (2016). Quantifying relative fish abundance with eDNA: a promising tool for fisheries management. *Journal of Applied Ecology*, 53(4), 1148-1157.

Larson, E. R., Graham, B. M., Achury, R., Coon, J. J., Daniels, M. K., Gambrell, D. K., ... & Suarez, A. V. (2020). From eDNA to citizen science: emerging tools for the early detection of invasive species. *Frontiers in Ecology and the Environment*, 18(4), 194-202.

Lear, G., Dickie, I., Banks, J., Boyer, S., Buckley, H. L., Buckley, T. R., ... & Holdaway, R. (2018). Methods for the extraction, storage, amplification and sequencing of DNA from environmental samples. *New Zealand Journal of Ecology*, 42(1), 10-50A.

Leese, F., Sander, M., Buchner, D., Elbrecht, V., Haase, P., & Zizka, V. M. (2021). Improved freshwater macroinvertebrate detection from environmental DNA through minimized nontarget amplification. *Environmental DNA*, 3(1), 261-276.

Leonard, J. A., Shanks, O., Hofreiter, M., Kreuz, E., Hodges, L., Ream, W., ... & Fleischer, R. C. (2007). Animal DNA in PCR reagents plagues ancient DNA research. *Journal of Archaeological Science*, 34(9), 1361-1366.

Leray, M., Yang, J. Y., Meyer, C. P., Mills, S. C., Agudelo, N., Ranwez, V., ... & 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), 1-14.

Leray, Matthieu, Joy Y. Yang, Christopher P. Meyer, Suzanne C. Mills, Natalia Agudelo, Vincent Ranwez, Joel T. Boehm, and Ryuji J. Machida. (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), 1-14.

Levi, T., Allen, J. M., Bell, D., Joyce, J., Russell, J. R., Tallmon, D. A., ... & Yu, D. W. (2019). Environmental DNA for the enumeration and management of Pacific salmon. *Molecular Ecology Resources*, 19(3), 597-608.

Levy-Booth, D. J., Campbell, R. G., Gulden, R. H., Hart, M. M., Powell, J. R., Klironomos, J. N., ... & Dunfield, K. E. (2007). Cycling of extracellular DNA in the soil environment. *Soil Biology and Biochemistry*, 39(12), 2977-2991.

Li, J., Lawson Handley, L. J., Read, D. S., & Hänfling, B. (2018). The effect of filtration method on the efficiency of environmental DNA capture and quantification via metabarcoding. *Molecular Ecology Resources*, 18(5), 1102-1114.

Lindahl, T. (1993). Instability and decay of the primary structure of DNA. *Nature*, 362(6422), 709-715.

Loeza-Quintana, T., Abbott, C. L., Heath, D. D., Bernatchez, L., & Hanner, R. H. (2020). Pathway to Increase Standards and Competency of eDNA Surveys (PISCeS)—Advancing collaboration and standardization efforts in the field of eDNA. *Environmental DNA*, 2(3), 255-260.

Longmire, J. L., Maltbie, M., & Baker, R. J. (1997). Use of" lysis buffer" in DNA isolation and its implication for Museum Collections. Museum of Texas Tech University.

Lopes, C. M., Sasso, T., Valentini, A., Dejean, T., Martins, M., Zamudio, K. R., & Haddad, C. F. B. (2017). eDNA metabarcoding: A promising method for anuran surveys in highly diverse tropical forests. *Molecular Ecology Resources*, 17(5), 904–914.

Lowe S., Browne M., Boudjelas S., De Poorter M. (2000) 100 of the World's Worst Invasive Alien Species A selection from the Global Invasive Species Database. Published by The Invasive Species Specialist Group (ISSG) a specialist group of the Species Survival Commission (SSC) of the World Conservation Union (IUCN), 12pp. First published as special lift-out in *Aliens* 12, December 2000. Updated and reprinted version: November 2004.

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.

Magoč, T., & Salzberg, S. L. (2011). FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics*, 27(21), 2957-2963.

Mahé, F., Rognes, T., Quince, C., de Vargas, C., & Dunthorn, M. (2014). Swarm: robust and fast clustering method for amplicon-based studies. *PeerJ*, 2, e593.

Malmström, H., Storå, J., Dalén, L., Holmlund, G., & Götherström, A. (2005). Extensive human DNA contamination in extracts from ancient dog bones and teeth. *Molecular Biology and Evolution*, 22(10), 2040-2047.

Martellini, A., Payment, P., & Villemur, R. (2005). Use of eukaryotic mitochondrial DNA to differentiate human, bovine, porcine and ovine sources in fecally contaminated surface water. *Water Research*, 39(4), 541-548.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal*, 17(1), 10-12.

Maruyama, A, Sugatani, K, Watanabe, K, Yamanaka, H, Imamura, A. (2018). Environmental DNA analysis as a non-invasive quantitative tool for reproductive migration of a threatened endemic fish in rivers. *Ecol Evol*. 8,11964– 11974.

Mathon, L., Valentini, A., Guérin, P. E., Normandeau, E., Noel, C., Lionnet, C., ... & Manel, S. (2021). Benchmarking bioinformatic tools for fast and accurate eDNA metabarcoding species identification. *Molecular Ecology Resources*.

Matsui, K., Honjo, M., & Kawabata, Z. (2001). Estimation of the fate of dissolved DNA in thermally stratified lake water from the stability of exogenous plasmid DNA. *Aquatic Microbial Ecology*, 26(1), 95-102.

McColl-Gausden, E. F., Weeks, A. R., Coleman, R. A., Robinson, K. L., Song, S., Raadik, T. A., & Tingley, R. (2021). Multispecies models reveal that eDNA metabarcoding is more sensitive than backpack electrofishing for conducting fish surveys in freshwater streams. *Molecular Ecology*, 30(13), 3111-3126.

McDevitt, A. D., Sales, N. G., Browett, S. S., Sparnenn, A. O., Mariani, S., Wangensteen, O. S., ... & Benvenuto, C. (2019). Environmental DNA metabarcoding as an effective and rapid tool for fish monitoring in canals. *Journal of Fish Biology*, 95(2), 679-682.

Menzel, P., Ng, K. L., & Krogh, A. (2016). Fast and sensitive taxonomic classification for metagenomics with Kaiju. *Nature Communications*, 7(1), 1-9.

Mercier, C., Boyer, F., Bonin, A., & Coissac, E. (2013). SUMATRA and SUMACLUST: fast and exact comparison and clustering of sequences. In Programs and Abstracts of the SeqBio 2013 workshop. Abstract (pp. 27-29).

Merkes, C. M., McCalla, S. G., Jensen, N. R., Gaikowski, M. P., & Amberg, J. J. (2014). Persistence of DNA in carcasses, slime and avian feces may affect interpretation of environmental DNA data. *PLoS One*, 9(11), e113346.

Minamoto, T., Miya, M., Sado, T., Seino, S., Doi, H., Kondoh, M., ... & Uchii, K. (2021). An illustrated manual for environmental DNA research: Water sampling guidelines and experimental protocols. *Environmental DNA*, 3(1), 8-13.

Miya, M., Gotoh, R. O., & Sado, T. (2020). MiFish metabarcoding: a high-throughput approach for simultaneous detection of multiple fish species from environmental DNA and other samples. *Fisheries Science*, 1-32.

Miya, M., Sato, Y., Fukunaga, T., Sado, T., Poulsen, J. Y., Sato, K., ... & Iwasaki, W. (2015). MiFish, a set of universal PCR primers for metabarcoding environmental DNA from fishes: detection of more than 230 subtropical marine species. *Royal Society Open Science*, 2(7), 150088.

Muha, T. P., Robinson, C. V., Garcia de Leaniz, C., & Consuegra, S. (2019). An optimised eDNA protocol for detecting fish in lentic and lotic freshwaters using a small water volume. *PloS one*, 14(7), e0219218.

Mychek-Londer, J. G., Balasingham, K. D., & Heath, D. D. (2020). Using environmental DNA metabarcoding to map invasive and native invertebrates in two Great Lakes tributaries. *Environmental DNA*, 2(3), 283-297.

Nguyen, B. N., Shen, E. W., Seemann, J., Correa, A. M., O'Donnell, J. L., Altieri, A. H., ... & Leray, M. (2020). Environmental DNA survey captures patterns of fish and invertebrate diversity across a tropical seascape. *Scientific Reports*, 10(1), 1-14.

Ogram, A., Sayler, G. S., & Barkay, T. (1987). The extraction and purification of microbial DNA from sediments. *Journal of Microbiological Methods*, 7(2-3), 57-66.

Olds, B. P., Jerde, C. L., Renshaw, M. A., Li, Y., Evans, N. T., Turner, C. R., ... & Lamberti, G. A. (2016). Estimating species richness using environmental DNA. *Ecology and Evolution*, 6(12), 4214-4226.

Parducci, L., Väliranta, M., Salonen, J. S., Ronkainen, T., Matetovici, I., Fontana, S. L., ... & Suyama, Y. (2015). Proxy comparison in ancient peat sediments: pollen, macrofossil and plant DNA. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 370(1660), 20130382.

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.

Philippot, L., Ritz, K., Pandard, P., Hallin, S., & Martin-Laurent, F. (2012). Standardisation of methods in soil microbiology: progress and challenges. *FEMS Microbiology Ecology*, 82(1), 1-10.

Piaggio, A. J., Engeman, R. M., Hopken, M. W., Humphrey, J. S., Keacher, K. L., Bruce, W. E., & Avery, M. L. (2014). Detecting an elusive invasive species: a diagnostic PCR to detect Burmese python in Florida waters and an assessment of persistence of environmental DNA. *Molecular ecology resources*, 14(2), 374-380.

Pietramellara, G., Ascher, J., Borgogni, F., Ceccherini, M. T., Guerri, G., & Nannipieri, P. (2009). Extracellular DNA in soil and sediment: fate and ecological relevance. *Biology and Fertility of Soils*, 45(3), 219-235.

Pilliod, D. S., Goldberg, C. S., Arkle, R. S., & Waits, L. P. (2013). Estimating occupancy and abundance of stream amphibians using environmental DNA from filtered water samples. *Canadian Journal of Fisheries and Aquatic Sciences*, 70(8), 1123-1130.

Pilliod, D. S., Goldberg, C. S., Arkle, R. S., & Waits, L. P. (2014). Factors influencing detection of eDNA from a stream-dwelling amphibian. *Molecular Ecology Resources*, 14(1), 109-116.

Pompanon, F., E. Coissac, and P. Taberlet. Metabarcoding, a new way of analysing biodiversity. *Biofutur* 319 (2011): 30-32.

Prodan, A., Tremaroli, V., Brolin, H., Zwinderman, A. H., Nieuwdorp, M., & Levin, E. (2020). Comparing bioinformatic pipelines for microbial 16S rRNA amplicon sequencing. *PLoS One*, 15(1), e0227434.

Quince, C., Lanzen, A., Davenport, R. J., & Turnbaugh, P. J. (2011). Removing noise from pyrosequenced amplicons. *BMC Bioinformatics*, 12(1), 1-18.

Ratsch, R., Kingsbury, B. A., & Jordan, M. A. (2020). Exploration of environmental DNA (eDNA) to detect Kirtland's snake (*Clonophis kirtlandii*). *Animals*, 10(6), 1057.

Rauschen, S., Schultheis, E., Pagel-Wieder, S., Schuphan, I., & Eber, S. (2009). Impact of Bt-corn MON88017 in comparison to three conventional lines on Trigonotylus caelestialium (Kirkaldy)(Heteroptera: Miridae) field densities. *Transgenic Research*, 18(2), 203-214.

Rees, H. C., Bishop, K., Middleditch, D. J., Patmore, J. R., Maddison, B. C., & Gough, K. C. (2014). The application of eDNA for monitoring of the Great Crested Newt in the UK. *Ecology and Evolution*, 4(21), 4023-4032.

Riascos, L., Geerts, A. N., Oña, T., Goethals, P., Cevallos-Cevallos, J., Berghe, W. V., ... & Boets, P. (2018). DNA-based monitoring of the alien invasive North American crayfish *Procambarus clarkii* in Andean lakes (Ecuador). *Limnologica*, 70, 20-25.

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

Schloss, P. D., Westcott, S. L., Ryabin, T., Hall, J. R., Hartmann, M., Hollister, E. B., ... & Weber, C. F. (2009). Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Applied and Environmental Microbiology*, 75(23), 7537-7541.

Schnell, I. B., Bohmann, K., & Gilbert, M. T. P. (2015). Tag jumps illuminated–reducing sequence-to-sample misidentifications in metabarcoding studies. *Molecular Ecology Resources*, 15(6), 1289-1303.

Schwentner M, Zahiri R, Yamamoto S, Husemann M, Kullmann B, et al. (2021) eDNA as a tool for non-invasive monitoring of the fauna of a turbid, well-mixed system, the Elbe estuary in Germany. *PloS One* 16(4): e0250452.

Scottish Fisheries Co-ordination Centre – SFCC (2007). Catch Fish Using Electrofishing Techniques. Introductory Electrofishing Training Manual. Fisheries Management SVQ Level 2, Inverness/Barony College, pp.44.

Sczyrba, A., Hofmann, P., Belmann, P., Koslicki, D., Janssen, S., Dröge, J., ... & McHardy, A. C. (2017). Critical assessment of metagenome interpretation—a benchmark of metagenomics software. *Nature Methods*, 14(11), 1063-1071.

Shaw, J. L., Clarke, L. J., Wedderburn, S. D., Barnes, T. C., Weyrich, L. S., & Cooper, A. (2016). Comparison of environmental DNA metabarcoding and conventional fish survey methods in a river system. *Biological Conservation*, 197, 131-138.

Shehzad, W., Riaz, T., Nawaz, M. A., Miquel, C., Poillot, C., Shah, S. A., ... & Taberlet, P. (2012). Carnivore diet analysis based on next-generation sequencing: Application to the leopard cat (Prionailurus bengalensis) in Pakistan. *Molecular Ecology*, 21(8), 1951-1965.

Shokralla, S., Spall, J. L., Gibson, J. F., & Hajibabaei, M. (2012). Next-generation sequencing technologies for environmental DNA research. *Molecular Ecology*, 21(8), 1794-1805.

Shu, L., Ludwig, A., & Peng, Z. (2020). Standards for methods utilizing environmental DNA for detection of fish species. *Genes*, 11(3), 296.

Shu, L., Ludwig, A., & Peng, Z. (2021). Environmental DNA metabarcoding primers for freshwater fish detection and quantification: In silico and in tanks. *Ecology and Evolution*.

Siegwald, L., Touzet, H., Lemoine, Y., Hot, D., Audebert, C., & Caboche, S. (2017). Assessment of common and emerging bioinformatics pipelines for targeted metagenomics. *PLoS One*, 12(1), e0169563.

Simmons, M., Tucker, A., Chadderton, W. L., Jerde, C. L., & Mahon, A. R. (2016). Active and passive environmental DNA surveillance of aquatic invasive species. *Canadian Journal of Fisheries and Aquatic Sciences*, 73(1), 76-83.

Singer, G. A. C., Fahner, N. A., Barnes, J. G., McCarthy, A., & Hajibabaei, M. (2019). Comprehensive biodiversity analysis via ultra-deep patterned flow cell technology: a case study of eDNA metabarcoding seawater. Scientific reports, 9(1), 1-12.

Sollazzo C., Scanu G., Aste F., Pineschi G., Belli M., Balzamo S., Martone C., Cadoni F., Bernabei S., et al. (2007). Protocollo di campionamento e analisi della fauna ittica dei sistemi lotici [Rapporto tecnico].

Spens, J., Evans, A. R., Halfmaerten, D., Knudsen, S. W., Sengupta, M. E., Mak, S. S., ... & Hellström, M. (2017). Comparison of capture and storage methods for aqueous macrobial eDNA using an optimized extraction protocol: advantage of enclosed filter. *Methods in Ecology and Evolution*, 8(5), 635-645.

Stoeckle, B. C., Beggel, S., Kuehn, R., & Geist, J. (2021). Influence of stream characteristics and population size on downstream transport of freshwater mollusk environmental DNA. *Freshwater Science*, 40(1), 191-201.

Strickland, G. J., & Roberts, J. H. (2019). Utility of eDNA and occupancy models for monitoring an endangered fish across diverse riverine habitats. *Hydrobiologia*, 826(1), 129-144.

Taberlet, P., Bonin, A., Zinger, L., & Coissac, E. (2018). Environmental DNA: For biodiversity research and monitoring. *Oxford University Press*.

Taberlet, P., Coissac, E., Hajibabaei, M. And Rieseberg, L.H. (2012), Environmental DNA. *Molecular Ecology*, 21: 1789-1793.

Takahara, T., Minamoto, T., & Doi, H. (2015). Effects of sample processing on the detection rate of environmental DNA from the Common Carp (*Cyprinus carpio*). *Biological Conservation*, 183, 64-69.

Takahara, T., Minamoto, T., Yamanaka, H., Doi, H., & Kawabata, Z. I. (2012). Estimation of fish biomass using environmental DNA. *PloS one*, 7(4), e35868.

Thomsen, P. F., Kielgast, J. O. S., Iversen, L. L., Wiuf, C., Rasmussen, M., Gilbert, M. T. P., ... & Willerslev, E. (2012). Monitoring endangered freshwater biodiversity using environmental DNA. *Molecular Ecology*, 21(11), 2565-2573.

Tordoni, E., Ametrano, C. G., Banchi, E., Ongaro, S., Pallavicini, A., Bacaro, G., & Muggia, L. (2021). Integrated eDNA metabarcoding and morphological analyses assess spatio-temporal patterns of airborne fungal spores. *Ecological Indicators*, 121, 107032.

Tréguier, A., Paillisson, J. M., Dejean, T., Valentini, A., Schlaepfer, M. A., & Roussel, J. M. (2014). Environmental DNA surveillance for invertebrate species: Advantages and technical limitations to detect invasive crayfish *Procambarus clarkii* in freshwater ponds. *Journal of Applied Ecology*, 51(4), 871-879.

Tucker, A. J., Chadderton, W. L., Jerde, C. L., Renshaw, M. A., Uy, K., Gantz, C., ... & Lodge, D. M. (2016). A sensitive environmental DNA (eDNA) assay leads to new insights on Ruffe (*Gymnocephalus cernua*) spread in North America. *Biological Invasions*, 18(11), 3205-3222.

Turner, C. R., Barnes, M. A., Xu, C. C., Jones, S. E., Jerde, C. L., & Lodge, D. M. (2014). Particle size distribution and optimal capture of aqueous macrobial eDNA. *Methods in Ecology and Evolution*, 5(7), 676-684.

Turner, C. R., Uy, K. L., & Everhart, R. C. (2015). Fish environmental DNA is more concentrated in aquatic sediments than surface water. *Biological Conservation*, 183, 93-102.

Valentini, A., Taberlet, P., Miaud, C., Civade, R., Herder, J., Thomsen, P. F., ... & Dejean, T. (2016). Next-generation monitoring of aquatic biodiversity using environmental DNA metabarcoding. *Molecular ecology*, 25(4), 929-942.

Vamos, E. E., Elbrecht, V., & Leese, F. (2017). Short COI markers for freshwater macroinvertebrate metabarcoding (No. e3037v2). *PeerJ* Preprints.

Varotto, C., Pindo, M., Bertoni, E., Casarotto, C., Camin, F., Girardi, M., ... & Cristofori, A. (2021). A pilot study of eDNA metabarcoding to estimate plant biodiversity by an alpine glacier core (Adamello glacier, North Italy). *Scientific reports*, 11(1), 1-13.

Vermeulen, S., and Koziell, I. (2002). Integrating global and local values: a review of biodiversity assessment. International Institute for Environment and Development, London, UK .

Vrijenhoek, R. (1994). DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Mol Mar Biol Biotechnol*, 3(5), 294-9.

Wang, S., Yan, Z., Hänfling, B., Zheng, X., Wang, P., Fan, J., & Li, J. (2021). Methodology of fish eDNA and its applications in ecology and environment. *Science of the Total Environment*, 755, 142622.

Wegleitner, B. J., Jerde, C. L., Tucker, A., Chadderton, W. L., & Mahon, A. R. (2015). Long duration, room temperature preservation of filtered eDNA samples. *Conservation Genetics Resources*, 7(4), 789-791.

Willerslev, E., Hansen, A. J., Binladen, J., Brand, T. B., Gilbert, M. T. P., Shapiro, B., ... & Cooper, A. (2003). Diverse plant and animal genetic records from Holocene and Pleistocene sediments. *Science*, 300(5620), 791-795.

Wright, E. S., Yilmaz, L. S., & Noguera, D. R. (2012). DECIPHER, a search-based approach to chimera identification for 16S rRNA sequences. *Applied and Environmental Microbiology*, 78(3), 717-725.

Wu, L., Wen, C., Qin, Y., Yin, H., Tu, Q., Van Nostrand, J. D., ... & Zhou, J. (2015). Phasing amplicon sequencing on Illumina Miseq for robust environmental microbial community analysis. *BMC Microbiology*, 15(1), 1-12.

Xia, Z., Zhan, A., Johansson, M. L., DeRoy, E., Haffner, G. D., & MacIsaac, H. J. (2021). Screening marker sensitivity: Optimizing eDNA-based rare species detection. *Diversity and Distributions*.

Xu, N., Zhu, B., Shi, F., Shao, K., Que, Y., Li, W., ... & Chang, J. (2018). Monitoring seasonal distribution of an endangered anadromous sturgeon in a large river using environmental DNA. *The Science of Nature*, 105(11), 1-7.

Yan, D., Mills, J. G., Gellie, N. J., Bissett, A., Lowe, A. J., & Breed, M. F. (2018). High-throughput eDNA monitoring of fungi to track functional recovery in ecological restoration. *Biological Conservation*, 217, 113-120.

Yoccoz, N. G., Bråthen, K. A., Gielly, L., Haile, J., Edwards, M. E., Goslar, T., ... & Taberlet, P. (2012). DNA from soil mirrors plant taxonomic and growth form diversity. *Molecular Ecology*, 21(15), 3647-3655.

Zhang, J., Kobert, K., Flouri, T., & Stamatakis, A. (2014). PEAR: a fast and accurate Illumina Paired-End reAd mergeR. *Bioinformatics,* 30(5), 614-620.

Zhu, B. (2006). Degradation of plasmid and plant DNA in water microcosms monitored by natural transformation and real-time polymerase chain reaction (PCR). *Water Research*, 40(17), 3231-3238.

Zimmermann, J., Jahn, R., & Gemeinholzer, B. (2011). Barcoding diatoms: evaluation of the V4 subregion on the 18S rRNA gene, including new primers and protocols. *Organisms Diversity & Evolution*, 11(3), 173-192.

Zinger, L., Bonin, A., Alsos, I. G., Bálint, M., Bik, H., Boyer, F., ... & Taberlet, P. (2019). DNA metabarcoding—Need for robust experimental designs to draw sound ecological conclusions.

5 SUPPLEMENTARY MATERIALS

Figure S1-S7 Fish species collected with the electrofishing technique.

Figure S1. *Alburnus alborella*

Figure S2. *Amiurus melas*

Figure S3. *Carassius* sp.

Figure S4. *Cyprinus carpio*

Figure S5. *Gambusia holbrooki*

Figure S6. *Lepomis gibbosus*

Figure S7. *Pseudorasbora parva*

Figure S8-S18. Specimens collected with the entomological net and photographed at the stereomicroscope with a Canon PowerShot G6 7.1MP Digital Camera. Zoom ranging from 0.32x to 3.2x. The sample ID is shown in top left corner of each photo. In parenthesis, the morphological identification / molecular identification are given (-: no molecular identification available).

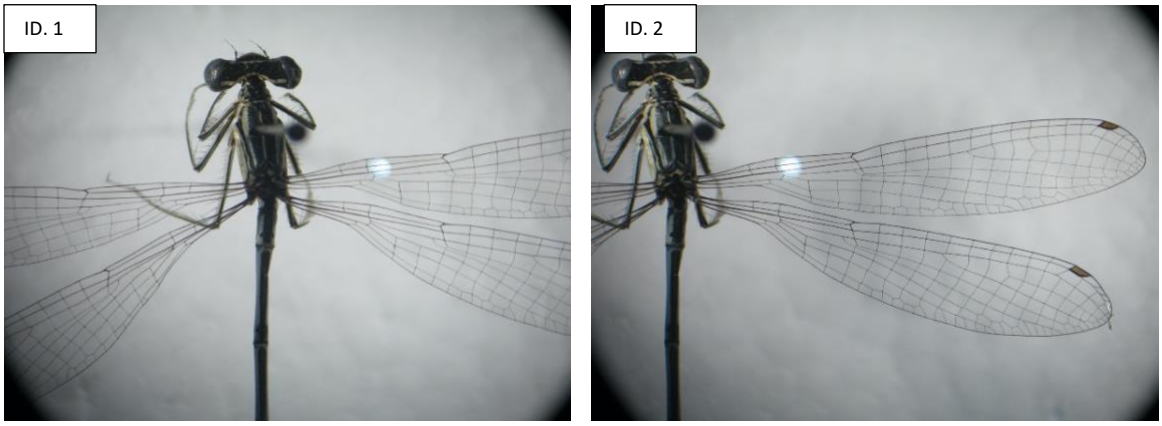
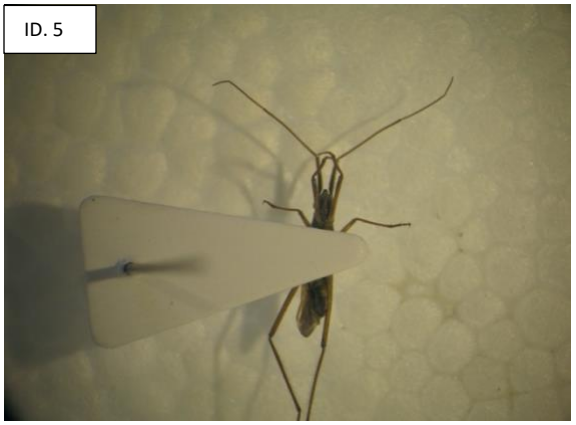


Figure S8. Dorsal vision. (*Platycnemis pennipes/Platycnemis pennipes*)

Figure S9. Dorsal vision.(Rynchota, Heteroptera/ -)

Figure S10. Dorsal vision. (Rynchota, Heteroptera / *Notostira elongata*)



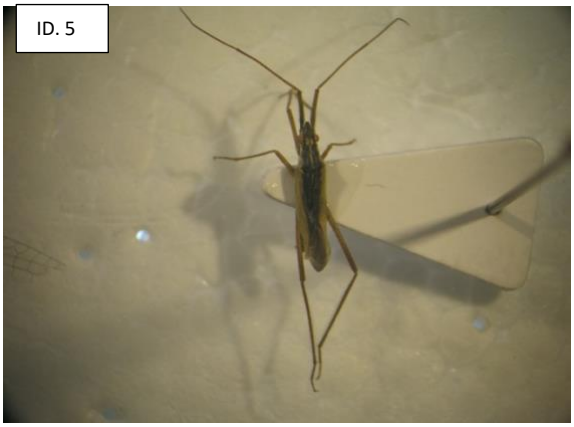


Figure S11. Ventral and dorsal vision (Rynchota, Heteroptera / *Notostira elongata*)

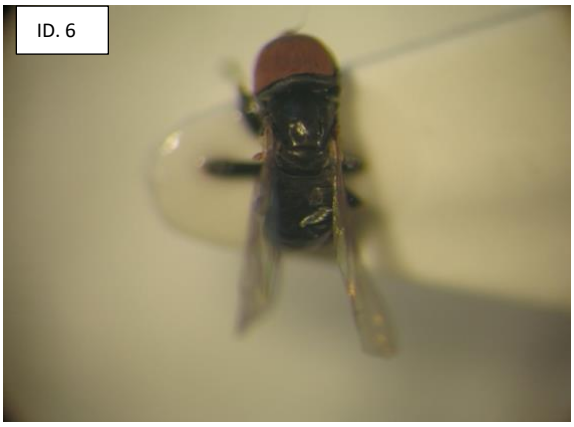


Figure S12. Lateral and dorsal vision. (Pipunculidae / -)

Figure S13. Dorsal and ventral vision. (Piesmidae / -)

Figure S14. Lateral vision. (Orthoptera / *Chorthippus dorsatus*)

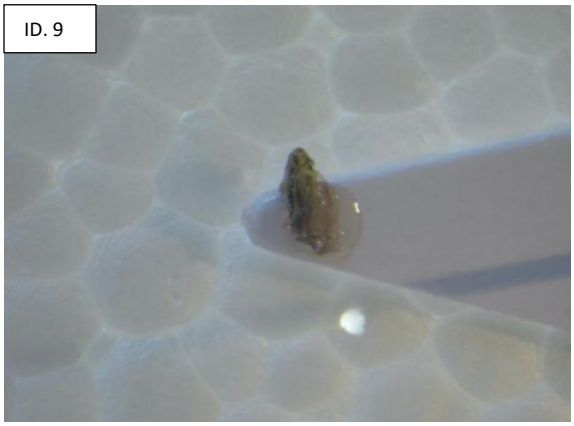


Figure S15. Lateral vision. (Rynchota, Homoptera / -)

Figure S16. Dorsal vision. (Rynchota / *Trigonotylus caelestialium*)

Figure S17. Dorsal and lateral vision. (Coleoptera / *Oedemera flavipes*)

Figure S18. Lateral vision. (Geometridae / *Peribatodes rhomboidaria*)

Figure S19-38. Specimens collected with the surber sampler and photographed at the stereomicroscope with a Canon PowerShot G6 7.1MP Digital Camera. Zoom ranging from 0.32x to 3.2x. The sample ID is shown in top left corner of each photo. . In parenthesis, the morphological identification / molecular identification are given (-: no molecular identification available).

Figure S19. Ventral and dorsal vision. (*Coenagrion* sp. / -)

Figure S20. Lateral vision. (*Coenagrion* sp. / -)

Figure S21. Dorsal vision. (*Valvata piscinalis* / -)

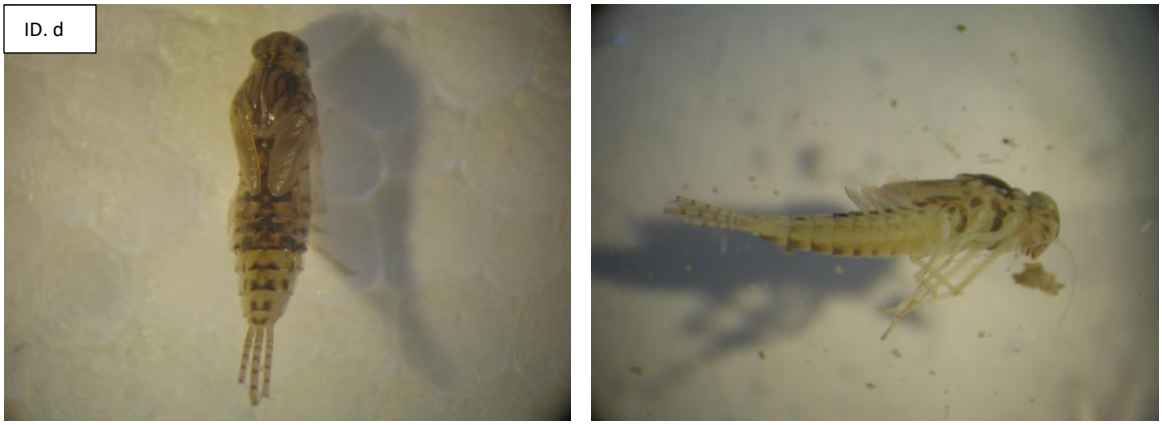


Figure S22. Dorsal and lateral vision. (*Procleon bifidum* / -)

Figure S23. Ventral vision. (*Physa fontinalis* / -)

Figure S24 Lateral, ventral and dorsal vision. (*Sigara* sp. / *Sigara lateralis*)

Figure S25. Lateral vision. (*Sigara* sp. / *Sigara lateralis*)

Figure S26. Lateral vision. (Tr. Chironomini / *Cryptochironomus obreptans*)

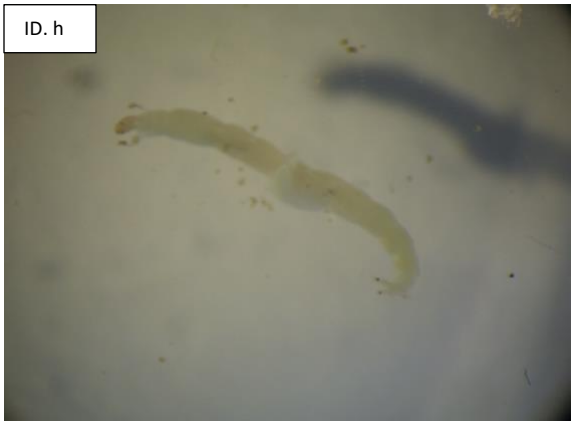


Figure S27. Lateral vision. (Chironomidae / *Parachironomus gracilior*)

Figure S28. Dorsolateral vision. (*Procleon bifidum / Cloeon dipterum*)

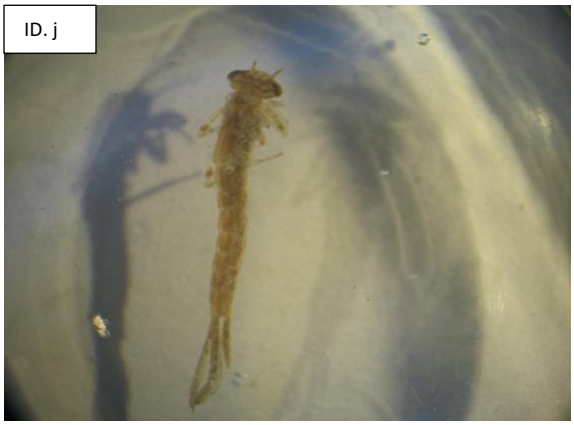


Figure S29. Dorsal vision. (*Coenagrion* sp./ -)

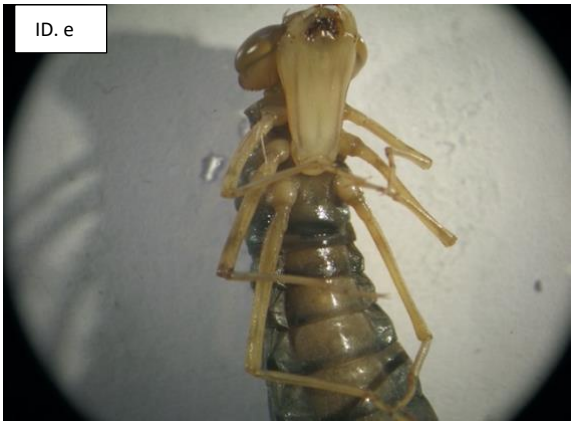


Figure S30. Dorsal and ventral vision. (*Anax imperator / Anax* sp. 1)

Figure S31. Lateral vision. (Orthocladinae / -)

Figure S32. Dorsal vision. (*Coenagrion* sp. / *Ischnura elegans*)

Figure S33. Dorsal vision. (*Coenagrion* sp. / *Ischnura elegans*)

Figure S34. Dorsal vision. (*Coenagrion* sp. / *Ischnura elegans*)

Figure S35. Dorsal vision. (*Coenagrion* sp. / *Ischnura elegans*)

Figure S36. Ventral vision. (*Coenagrion* sp. / *Ischnura elegans*)

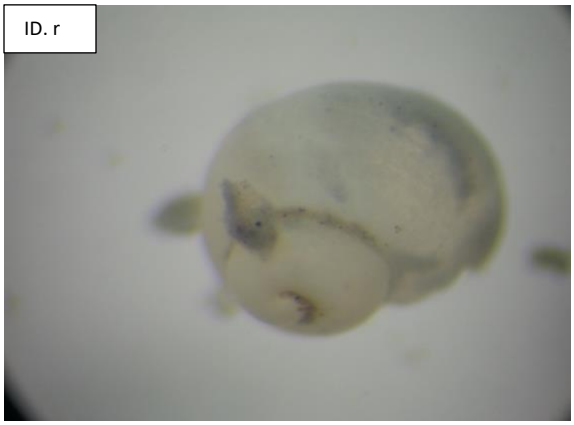


Figure S37. Dorsolateral vision. (Valvata / -)

Figure S38. (Chironomidae (eggs) / -)