Methods

A DNA metabarcoding protocol for hyporheic freshwater meiofauna: Evaluating highly degenerate COI primers and replication strategy

Alexander M. Weigand[1,2], **Jan-Niklas Macher**[1,3]

1 *University of Duisburg-Essen, Universitätsstraße 5, 45141 Essen, Germany.*

2 *Musée National d'Histoire Naturelle, 25 Rue Münster, 2160 Luxembourg, Luxembourg.*

3 *Naturalis Biodiversity Center, Pesthuislaan 7, 2333 BA Leiden, The Netherlands.*

Corresponding author: *Alexander M. Weigand* (alexander.weigand@uni-due.de)

Abstract

The hyporheic zone, i.e. the ecotone between surface water and the groundwater, is a rarely studied freshwater ecosystem. Hyporheic taxa are often meiofaunal (<1 mm) in size and difficult to identify based on morphology. Metabarcoding approaches are promising for the study of these environments and taxa, but it is yet unclear if commonly applied metabarcoding primers and replication strategies can be used. In this study, we took sediment cores from two near natural upstream (NNU) and two ecologically improved downstream (EID) sites in the Boye catchment (Emscher River, Germany), metabarcoding their meiofaunal communities. We evaluated the usability of a commonly used, highly degenerate COI primer pair (BF2/BR2) and tested how sequencing three PCR replicates per sample and removing MOTUs present in only one out of three replicates impacts the inferred community composition. A total of 22,514 MOTUs were detected, of which only 263 were identified as Metazoa. Our results highlight the gaps in reference databases for meiofaunal taxa and the potential problems of using highly degenerate primers for studying samples containing a high number of non-metazoan taxa. Alpha diversity was higher in EID sites and showed higher community similarity when compared to NNU sites. Beta diversity analyses showed that removing MOTUs detected in only one out of three replicates per site greatly increased community similarity in samples. Sequencing three sample replicates and removing rare MOTUs is seen as a good compromise between retaining too many false-positives and introducing too many false-negatives. We conclude that metabarcoding hyporheic communities using highly degenerate COI primers can provide valuable first insights into the diversity of these ecosystems and highlight some potential application scenarios.

Key Words

Hyporheos DNA Metabarcoding, Streambed Health Assessment, Biological Quality Element, Meiofaunal Enrichment

Introduction

DNA barcoding has become a standard technique in ecological and biodiversity studies (Hebert et al. 2003, Macher et al. 2016, Valentini et al. 2009). The technique has been enhanced by the possibility to generate and analyse millions of sequences on high-throughput sequencing machines. The use of DNA metabarcoding (Taberlet et al. 2012) is especially beneficial for studies on whole communities and taxa that are difficult to iden-tify based on their morphology. DNA metabarcoding has been tested and is used for the study of various habitats and different taxonomic groups (Arribas et al. 2016, Gibson et al. 2015, Yu et al. 2012) and recent studies have shown that metabarcoding can be used for standardised bioassessments in aquatic ecosystems (Cordier et al. 2017, Elbrecht et al. 2017, Vasselon et al. 2017). In this context, a recently formed consortium (EU COST Action CA15219 '*DNAqua-Net*') comprising researchers, stakeholders, water managers and entrepreneurs has the

goal to streamline the use of molecular methods for water quality assessment and to include them in international programmes such as the EU Water Framework Directive and EU Marine Strategy Framework Directive (Leese et al. 2016, 2018).

To date, the assessment of stream biodiversity and quality often relies on the study of presence, absence and abundance of diatoms, macrophytes, fish and macrozoobenthic invertebrates (MZB), i.e. invertebrates living on or in the riverbed (Birk et al. 2012). While this approach has been shown to be effective for biomonitoring, it neglects the fact that streams and rivers are complex ecosystems harbouring a wide range of microhabitats and taxonomic groups. Several faunistic groups such as the hyporheic community inhabiting the soil below the riverbed, i.e. the ecotone between the riverine surface water and the shallow groundwater, are largely neglected in current stream ecosystem biomonitoring. Hyporheic habitats harbour great biodiversity, have important ecological functions and thereby also provide crucial ecosystem services: They serve as i) a buffer for stressors in the water by filtering, storing and thus removing stressors from the water column, ii) a retreat for animals during droughts and other unfavourable environmental conditions at the surface, leading to a higher resilience of freshwater communities, iii) a spawning area for fish and invertebrates and iv) a nursery habitat for larvae (Datry et al. 2017, Lawrence et al. 2013, Stubbington et al. 2016). Further, hyporheic communities are known to react sensitively to stressors such as pollution, while also showing high pioneer settlement rates in newly available habitats (Beier and Traunspurger 2001, Eisendle-Flöckner et al. 2013). This means that insights into the biodiversity and ecology of hyporheic communities can potentially complement current stream quality assessments by providing information on the neglected biodiversity in the streambed and surface-subsurface hydrological connections (Lawrence et al. 2013).

So far, the hyporheic freshwater fauna has been scarcely studied mainly due to difficulties in sampling (mostly pumping and digging; Fraser and Williams 1997) and the small size of many taxa, making sample sorting and subsequent species identification a challenge (Giere 2008, Hakenkamp and Palmer 2000). As species identification has been shown to be often inaccurate even for relatively well known macrozoobenthic taxa (Haase et al. 2010), this inaccuracy is expected to be even more severe for the often difficult to identify meiofaunal (i.e. <1 mm body length) taxa inhabiting the hyporheic zone. However, this problem can be tackled with metabarcoding approaches, which often allow identification of taxa when genetic information has been deposited in a reference database such as the Barcode of Life Data System (BOLD, Ratnasingham and Hebert 2007). Studies on marine ecosystems have shown that molecular approaches can be successfully applied to identify soil fauna (Creer et al. 2010, Fonseca et al. 2014, Guardiola et al. 2015).

We here present and evaluate a DNA metabarcoding protocol for the assessment of hyporheic freshwater communities, targeting the meiofauna (i.e. organisms <1 mm but >42 μm) of streambed sediments (i.e. the hyporheic zone). Further, we test the performance of a highly degenerate primer set, amplifying a fragment of the mitochondrial cytochrome c oxidase subunit I (COI) gene for the assessment of hyporheic communities and investigate the effect of PCR replication strategy on the inferred biodiversity parameters and community composition.

Testing primer performance and replication strategy is of high importance, as even though it is widely accepted that the mitochondrial COI marker is the gene of choice for metabarcoding of freshwater invertebrates, no consensus has yet been reached regarding primer sets and the number of PCR replicates needed for reliable detection of taxa. Further, to the best of our knowledge, no studies have tested the performance of highly degenerate COI primers for metabarcoding hyporheic freshwater communities and replication strategy for studying these habitats has also not been tested.

Different primers have been used and evaluated for metabarcoding of freshwater invertebrates (Andújar et al. 2018, Blackman et al. 2017, Elbrecht and Leese 2017, Vamos et al. 2017) and the BF2/BR2 combination introduced by Elbrecht et al. (2017) has been found to reliably amplify a wide range of taxonomic groups, including taxa typically inhabiting the hyporheic zone, such as mites, dipteran larvae and oligochaetes. One prominent exception is the Nematoda, which is often not picked up by COI primers despite being abundant in soil habitats (Creer et al. 2010, Avó et al. 2017). However, recent studies have found that highly degenerate COI primers can amplify a great number of non-metazoan taxa when applied to samples containing a high amount of such taxa (Haene et al. 2017, Macher and Leese 2017, Wangensteen et al. 2018). As this is usually the case in sediment samples taken from the hyporheic zone, testing the performance of highly degenerate primers is an urgent task for planning further studies.

Different replication strategies for metabarcoding studies have been tested (Alberdi et al. 2017, Ficetola et al. 2015) and some studies argue that sequencing depth and sample number can be more important than replication (Ficetola et al. 2015, Smith and Peay 2014). Sequencing PCR replicates allows for the identification and subsequent removal of potential false-positives by retaining only taxa that are found in multiple replicates per sample. However, replicates can fail to sequence or can be influenced by stochastic effects such as primer and PCR bias as well as random sample picking during sequencing (Alberdi et al. 2017). These effects can affect the abundance of specific PCR templates and (subsequent) sequencing reads and can hence lead to false-negatives due to the detection of taxa in one replicate, but not in the other. Therefore, sequencing more than two replicates and considering only taxa present in the majority of replicates might increase the robustness of assessment results (Alberdi et al. 2017). Testing replication strategy when applying highly degenerate COI primers to hyporheic

communities is important, as the expected large number of amplified non-metazoan taxa might impact results due to their competition for sequencing reads with the metazoan target taxa.

We here present and evaluate a DNA metabarcoding protocol for the assessment of meiofaunal hyporheic freshwater communities. Further, we test the performance of a highly degenerate primer set amplifying a fragment of the mitochondrial cytochrome c oxidase subunit I (COI) gene for the assessment of hyporheic communities and investigate the effect of PCR replication strategy on the inferred community composition. The here presented metabarcoding protocol is further discussed in the context of potential application scenarios.

Material and methods

Sampling procedure

Sediment was collected at five lotic freshwater sites in the Boye tributary of the Emscher River catchment (Ruhr Metropolitan Region, North Rhine-Westphalia, Germany), which is one of Europe's largest river management and restoration projects (Perini 2016, Semrau and Hurck 2012). The sampling sites can be classified into two categories: near-natural upstream regions (NNU; three sites: NNU_58, 61, 101) and ecologically-improved downstream regions with connection to a near-natural upstream section (EID: two sites: EID_27, 57) (Coordinates: Suppl. material 4: Table S1; see also Winking et al. 2014). The two EID-sites represent ecologically-improved sewage channels. All five localities were visited on 15 October 2015. The NNU-site 'Scheidgensbach' (NNU_61) was found to be water-depleted, but was still sampled to investigate the applicability of our DNA-based bioassessment method for intermittent freshwaters. Each stream was sampled by taking nine subsamples, which were taken in different microhabitats over an area of 10 metres to account for fine-scale patterns in local hyporheic community structures. Each subsampling was performed by using a disposable plastic syringe with removed tip as the corer (diameter: 2.8 cm; height: 7.5 cm; volume: 50 mL). The syringe was pushed into the sediment without the plunger until the 50 mL label was below the sediment surface. The plunger was then placed on the syringe, thus creating an underinflation in the syringe. The now closed syringe was pulled up and the collected sediment immediately stored in three times the amount of 96% ethanol. The uppermost 2 cm of sediment were discarded to prevent sampling of benthic taxa. This procedure was repeated nine times, sampling the different microhabitats present at a site. A new syringe was used for each sampling site. All sediment cores taken at a given site were pooled. Sediment at the sampling sites was mostly very fine and contained high amounts of organic material (Suppl. material 2: Figure S1).

Extraction of hyporheic meiofauna from sediment and DNA isolation

To extract hyporheic meiofaunal organisms from the sediment samples, the sediment was sieved through 1 mm mesh size in the laboratory, using high-pressure water to flush sediment through sieves. Similar sieving methods have been tested before (Wangensteen and Turon 2017). The >1 mm fraction was collected and stored in 96% ethanol, but was not further processed. The <1 mm fraction was sieved through a second sieve with 42 µm mesh size. This step was repeated multiple times, as some samples contained a high amount of fine sediments leading to fast clogging of the 42 µm sieve. The <42 µm fraction was discarded, whereas the >42 µm (and <1 mm) fraction was stored in 96% ethanol and used for subsequent DNA extraction. When a sample was split into several sub-samples for sieving, the retained material from all sub-samples was pooled and vortexed. All sieves were sterilised by bleaching overnight before processing a new sample.

The following protocol for genomic DNA extraction from lotic freshwater sediments is available as a step-by-step laboratory handout (Suppl. material 1). For each sample, ethanol was removed by evaporation over a period of three days at room temperature (RT) under a sterile flow hood. The dried sediment was weighed and, if a sample contained >10 g sediment, it was fractionated and placed into several Petri dishes. Cell lysis was achieved by adding 25 mL of a TNES-Proteinase K buffer (39:1 parts by volume; 3000 U/mL Proteinase K) to each Petri dish containing a maximum of 10 g sediment. Sediment and lysis buffer were homogeneously mixed by multiple pipetting. The dish was covered with a lid and closed with parafilm. Samples were incubated for ~5 h at RT and were mixed every hour by pipetting. The solution was transferred into a Falcon-tube and centrifuged for 1 min at 2500 xg (RT). Supernatant was transferred into a new Falcon-tube and used for DNA extraction via a salt precipitation protocol (Sunnucks and Hales 1996). Per 1 mL solution, 350 µL 5M NaCl was added and the sample was vortexed. Then, the sample was centrifuged for 5 minutes at 2500 xg (RT) and supernatant was transferred into another Falcon-tube. Per 1 mL solution, 1 mL of ice-cold 100% ethanol was added. In case of higher volumes, the sample was split into subsamples. Each (sub)sample was centrifuged for 15 min at 3000 xg at 4 °C. Supernatant was discarded and 4 mL of 70% ethanol was added to the pellet. Each (sub)sample was again centrifuged for 15 min at 3000 xg at 4 °C. Ethanol was removed and the pellet air-dried by evaporation. Pellets were resuspended in 10 mL lab grade water. When samples were split into subsamples, the very same 10 mL were used for resuspension of all pellets belonging to one sample, effectively pooling all subsamples and keeping the final volume to a minimum. The PowerMax DNA Isolation Kit (MO BIO Laboratories, QIAGEN, Germany) was used for the final extraction of DNA from sediments and for removal of PCR inhibitors.

We also tried to extract DNA according to the standard PowerMax protocol, i.e. without the additional steps described above. However, this resulted in minimal DNA yield and we thus developed the protocol described here.

The resuspended pellets (10 mL volume) were transferred into the provided BeadTubes and extraction was carried out as described in the kit manual, with a final elution volume of 5 mL. The eluted DNA was finally concentrated by another salt precipitation as described above and the DNA pellet was resuspended in 50 µL PCR grade water.

COI amplification and sequencing

Extracted DNA was quantified on a Fragment Analyzer with the Standard Sensitivity Genomic Kit (Advanced Analytical, Oak Tree, USA) and a volume containing 15 ng of DNA was used for PCR. A two-step PCR protocol was followed: For the first PCR, untailed primers BF2 and BR2 (Elbrecht and Leese 2017) were used for amplification of a 421 bp fragment of COI using Illustra PuReTaq Ready-to-go PCR beads (GE Healthcare, Little Chalfont, UK). Initial denaturation was performed for 3 minutes at 94 °C, followed by 25 cycles at 94 °C for 30 seconds, 50 °C for 30 seconds, 72 °C for two minutes, followed by a final elongation at 72 °C for five minutes. For the second PCR, 1 µL of the product was used with individually tagged BF2 / BR2 primers (combinations: Suppl. material 5: Table S2). The PCR was conducted as described above, but limited to 15 cycles. Three independent PCRs were run for each sample. A negative control was included in all PCR steps and DNA concentrations in negative controls were checked on gels and the Fragment Analyzer using the NGS High Sensitivity Kit. As no DNA was observed in the negative controls, these were not sequenced. PCR products were cleaned using the MinElute PCR Purification Kit (QIAGEN, Germany), quantified on the Fragment Analyzer using the NGS High Sensitivity Kit, size selected using SpriSelect beads (left sided size selection, ratio 0.74×; Beckman Coulter, Brea, USA) and equimolarly pooled before being sent for sequencing on one run of the Illumina MiSeq platform (v2 Kit, 2× 250bp reads) at GATC Biotech (Constance, Germany).

Data quality filtering and MOTU clustering

All raw reads were processed with R scripts as in Elbrecht et al. (2016). Reads were demultiplexed and paired-end reads were merged using USEARCH (v.8.1.1756; Edgar 2010) with the maximum error rate set to fastq_maxee = 1. Primers were removed with cutadapt (v.1.9, Martin 2011). Sequences were dereplicated, singletons were removed and the UPARSE pipeline (Edgar 2013) was used for clustering at 97% similarity, which is a common approach in invertebrate metabarcoding studies relying on the barcoding gap approach (Bista et al. 2017, Elbrecht et al. 2017). Subsequently, only Molecular Operational Taxonomic Units (MOTUs) that had read abundances over 0.001% per replicate (corresponding to >1 read in replicates with >100,000 reads) were retained, while all other MOTUs were discarded. This step has been shown to be a suitable alternative to rarefaction (Elbrecht and Leese 2015).

Taxonomic assignment of MOTUs

MEGAN (v.6.8.18, Huson et al. 2007) was used to assign Linnaean taxonomy to the generated MOTUs by comparing sequences against a custom-made database containing all dereplicated COI sequences of <5000 bp length available from the NCBI GenBank (Benson et al. 2017) (GenBank sequences downloaded on 10 September 2017; MEGAN settings: -evalue 1e-60, -max_target_seqs 10). The following thresholds were used for identification: species (97% identity), genus (95%), family (93%), order (90%) and kingdom (85%). Metazoan MOTUs were then further checked against the Barcode of Life Data System (BOLD, Ratnasingham and Hebert 2007). In case of incongruent taxonomic classifications, the highest shared taxonomic level was used (e.g. family if different genera within the same family were proposed). MOTUs identified as "Metazoa", but referring to "Invertebrate Environmental Samples" tagged with "Eukaryota; Metazoa; environmental samples" in NCBI and MOTUs with strongly conflicting taxonomic assignments in NCBI and BOLD (e.g. identified as fungi or bacteria in one database, but as Metazoa in the other) were excluded from the final list of metazoan MOTUs, as were five pseudogenes of the *Homo sapiens* COI gene (pseudogene MTCO1P8, 13, 28, 49 and 51 with ≥99% sequence identity each). The taxonomy of oligochaete MOTUs was further manually resolved using the recently published Swiss database of Vivien et al. (2017).

Diversity analyses

Metazoan species richness was determined using the R package vegan (Oksanen et al. 2007), which was also applied to calculate beta diversity (Sørensen dissimilarity) across sampling sites and between replicates within single sampling sites. All calculations were done for Metazoa and the two most abundant taxonomic groups within the metazoan dataset, i.e. Oligochaeta and Diptera. Analyses were run on two datasets: first for the dataset including all MOTUs found per sampling site and second, for the dataset including only MOTUs found in at least two out of three replicates. All diversity estimates were calculated excluding the water-depleted site NNU_61. The Venn diagram creator of the University of Ghent (http=//bioinformatics.psb.ugent.be/webtools/Venn/) was used to visualise the number of metazoan MOTUs shared between sampling sites and number of MOTUs shared between replicates within sampling sites.

Results

No DNA was observed in the negative controls processed together with samples and thus, no contamination was suspected. A total of 14,128,336 read pairs were ob-
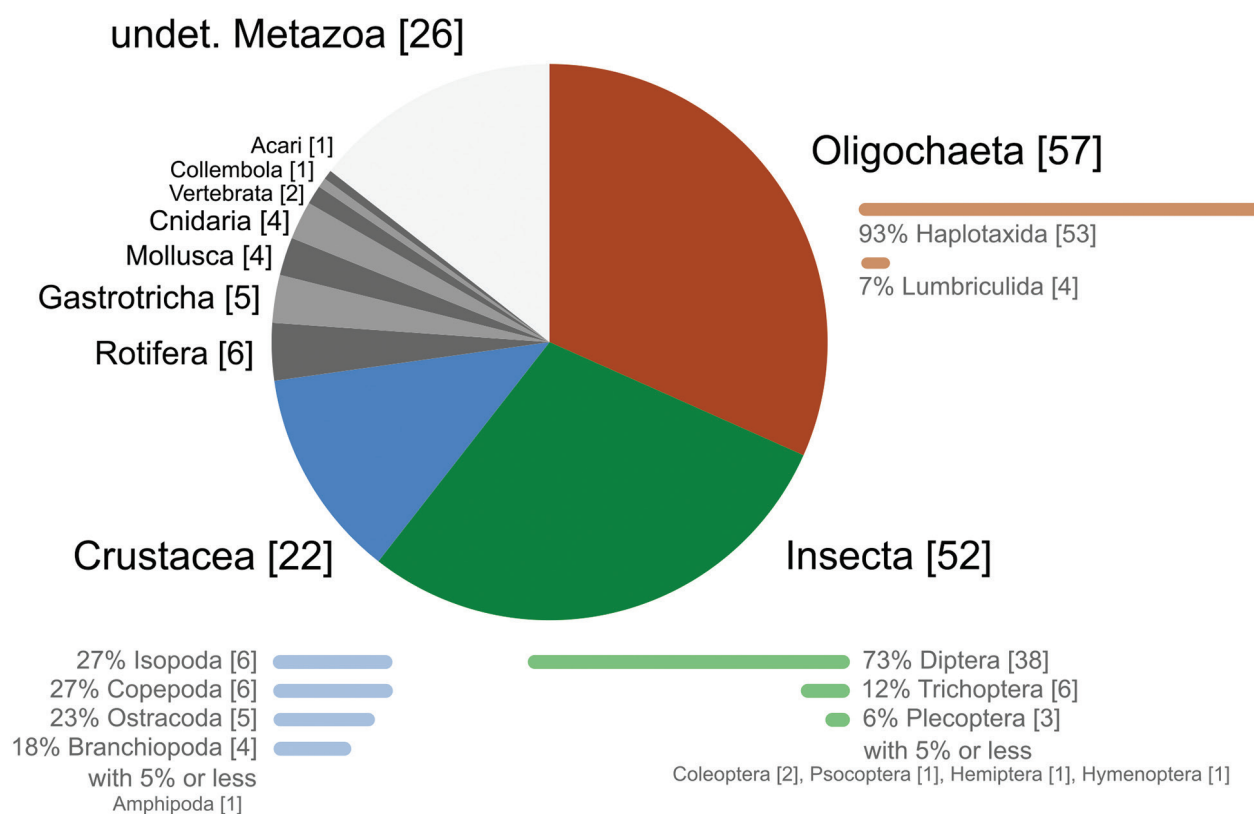
undet. Metazoa [26]

Acari [1]
Collembola [1]
Vertebrata [2]
Cnidaria [4]
Mollusca [4]
Gastrotricha [5]
Rotifera [6]

Oligochaeta [57]

93% Haplotaxida [53]

7% Lumbriculida [4]

Crustacea [22]

Insecta [52]

27% Isopoda [6]
27% Copepoda [6]
23% Ostracoda [5]
18% Branchiopoda [4]
with 5% or less
Amphipoda [1]

73% Diptera [38]
12% Trichoptera [6]
6% Plecoptera [3]
with 5% or less
Coleoptera [2], Psocoptera [1], Hemiptera [1], Hymenoptera [1]

Figure 1. Taxonomic composition of hyporheic community. Taxonomic composition of the hyporheic community revealed for the filtered dataset with the 2-out-of-3 replicate strategy applied, excluding the water-depleted site NNU_61.

tained after sequencing, retaining 2,313,243 high quality reads after bioinformatic processing and quality filtering (Suppl. material 6: Table S3).

Taxonomic composition

A total of 22,514 MOTUs were recovered after bioinformatic processing. Of these, 3,935 (17.5%) could be assigned to a taxonomic name on Kingdom level with the majority being identified as bacteria (2,806; 71.3%). After applying the taxonomic assignment procedure described above, a total of 263 metazoan MOTUs were retained across all five sampling sites (Suppl. material 2: Figure S2, Suppl. material 7: Table S4), resulting in 33.6% of all reads belonging to metazoan taxa. We further reduced the final dataset by a) excluding the water depleted site NNU_61, for which only two PCR replicates were available and b) by retaining only MOTUs found in at least two out of three of the sampling site replicates. Thus, the dataset was further reduced to 180 metazoan MOTUs, represented by Oligochaeta (n = 57), Insecta (n = 52; mainly Diptera [38] and Trichoptera [6]), Crustacea (22; mainly Isopoda [6], Copepoda [6], Ostracoda [5] and Branchiopoda [4]), Rotifera (6), Gastrotricha (5), Mollusca (4), Cnidaria (4), Collembola (1), vertebrates (2) and Acari (1) (Table 1, Figure 1, taxa list: Suppl. material 8: Table S5). Taxonomic resolution was moderate, with 61% of metazoan MOTUs (109/180)

Table 1. Overview of taxonomic resolution for metazoan MOTUs. Number of Metazoa, Diptera and Oligochaeta MOTUs identified per taxonomic level when only MOTUs present in at least two out of three replicates per sampling site were considered.

MOTUs	Kingdom level	Order level	Family level	Genus level	Species level
Metazoa	180	154	145	124	109
Oligochaeta		57	53	47	43
Diptera		38	37	33	28

identified up to the species level, 69% (124/180) at least to genus and 81% (145/180) at least to family level. For the two most MOTU-rich taxa, taxonomic resolution was considerably higher: Oligochaeta (species-level: 75%; genus: 83%; family: 93%), respectively, Diptera (species: 74%; genus: 87%; family: 97%) (Figure 2). Twenty six MOTUs (14%) were only identified to the level of Metazoa.

Alpha diversity

Alpha diversity of all metazoan taxa was highest in the ecologically improved downstream (EID) sites, with 150 MOTUs found in EID_57 and 134 MOTUs in EID_27 when analysing the full dataset, i.e. all MOTUs found at a sampling site, respectively, 99 metazoan MOTUs in EID_57 and 68 MOTUs in EID_27, when only MOTUs
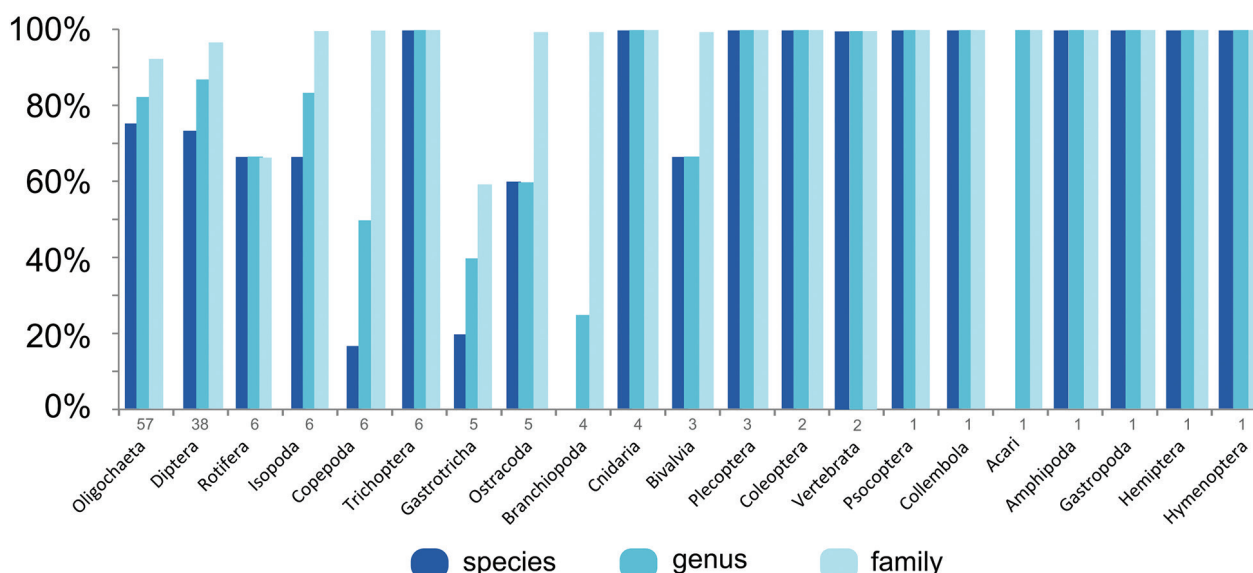
Figure 2. Taxonomic resolution of hyporheic community. Taxonomic resolution of the 2-out-of-3 replicate dataset, excluding the water-depleted site NNU_61. Numbers on the x-axis indicate the number of MOTUs per taxonomic group.

Table 2. Overview of alpha diversity estimates. Alpha diversity of Metazoa, Oligochaeta and Diptera MOTUs for the complete dataset including all MOTUs per site, respectively, the dataset excluding MOTUs found in only one out of three PCR replicates.

Taxa	EID_57	EID_27	NNU_58	NNU_101
All metazoan MOTUs	150	134	65	82
Metazoan MOTUs in 2 out of 3 replicates	99	68	44	44
All Oligochaeta MOTUs	52	53	11	30
Oligochaeta MOTUs in 2 out of 3 replicates	34	35	7	14
All Diptera MOTUs	25	24	19	16
Diptera MOTUs in 2 out of 3 replicates	17	13	15	9

Table 3. Overview of beta diversity estimates. Sørensen dissimilarity between replicates within sampling sites for all Metazoa, Oligochaeta and Diptera MOTUs and Sørensen dissimilarity calculated across sampling sites.

All metazoan MOTUs	All 3 replicates	2-out-of-3 replicates
Site	**Sørensen dissimilarity**	
EID_27	0.31	0.09
EID_57	0.22	0.08
NNU_58	0.23	0.10
NNU_101	0.29	0.07
Across sampling sites	0.60	0.75
Oligochaeta MOTUs		
EID_27	0.18	0.02
EID_57	0.17	0.02
NNU_58	0.26	0.11
NNU_101	0.29	0.05
Across sampling sites	0.52	0.71
Diptera MOTUs		
EID_27	0.31	0.11
EID_57	0.21	0.09
NNU_58	0.15	0.07
NNU_101	0.28	0.08
Across sampling sites	0.64	0.77

found in two out of three replicates were analysed. Both near natural upstream (NNU) sites demonstrated a considerably lower number of MOTUs than the EID sites (full dataset: 82 for NNU_101, respectively, 65 for NNU_58; 2-out-of-3 dataset: both 44 MOTUs) (Figure 3). The same diversity patterns were found for Oligochaeta and Diptera MOTUs (Figure 4, all values: Table 2).

Beta diversity

When all MOTUs found at a sampling site were analysed, Sørensen dissimilarity within sampling sites ranged from 0.15 (Diptera, site NNU_58) to 0.31 (Metazoa and Diptera, site 27). In contrast, when only MOTUs found in at least two of three replicates per site were included in the analyses, communities in replicates were much more similar, with Sørensen dissimilarity ranging from 0.02 (Oligochaeta, sites 27 and 57) to 0.11 (Oligochaeta, site 58 and Diptera, site 27) (all values: Table 3).

Beta diversity (Sørensen dissimilarity) across all sampling sites was 0.60 (Metazoa), 0.52 (Oligochaeta) and 0.64 (Diptera) when MOTUs from all three replicates per site were analysed and increased when only MOTUs present in at least two replicates per site were analysed: 0.75 (Metazoa), 0.71 (Oligochaeta) and 0.77 (Diptera) (Table 3).

Metazoan communities were found to be more similar between the two ecologically improved downstream (EID) sites (Sørensen dissimilarity: 0.60) than between
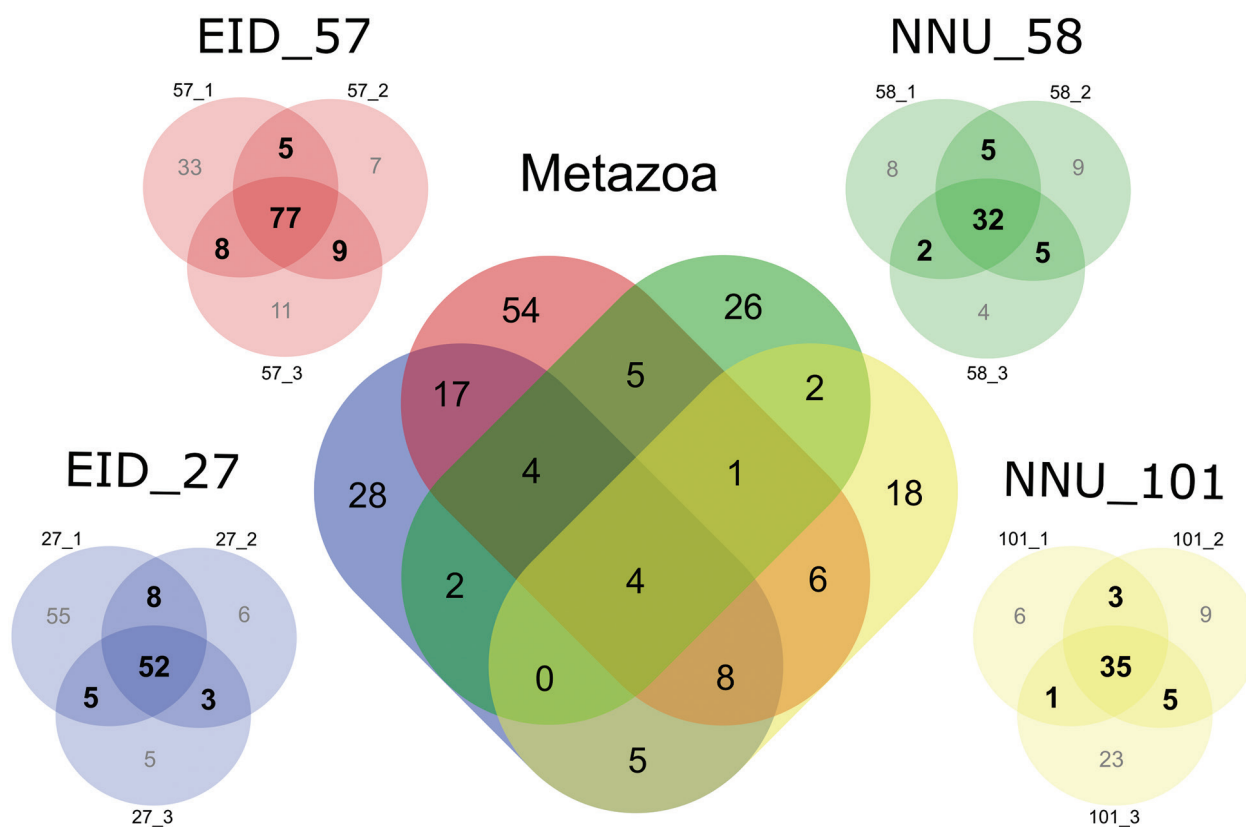
Figure 3. Overlap of MOTUs between PCR replicates and between sites for all metazoan taxa. Venn diagrams of metazoan MOTUs found per PCR replicate, number of MOTUs shared between replicates and number of MOTUs shared between sampling sites when only MOTUs present in at least two of three replicates per sampling site were considered.

the two near natural upstream (NNU) sites (Sørensen dissimilarity: 0.84). The NNU sites shared only seven MOTUs, all of which also occurred at EID sites: *Limnodrilus hoffmeisteri* lineage VII (MOTU_4), *Micropsectra notescens* (MOTU_10), two MOTUs from the *Hydra vulgaris* complex (MOTU_160, MOTU_2580), *Homo sapiens* (MOTU_186), a Chaetonotidae (MOTU_592) and an unidentified metazoan (MOTU_1959). In contrast, the EID sites shared 33 MOTUs primarily comprising Oligochaeta (18/33; 55%) and Diptera (5/33; 15%), with 17 of these exclusively occurring at EID sites.

Discussion

Degenerate COI primers for metabarcoding of hyporheic communities

The applied metabarcoding approach detected taxa characteristic for the hyporheic community (Pacioglu 2010), with water mites, platyhelminthes and nematodes comprising prominent exceptions. However, these taxa were observed in preliminary morphological studies performed in the course of two B.Sc. theses at the same sites and using the same sampling technique, seven weeks prior to DNA metabarcoding sampling (Frie 2016, Tonscheidt 2016) (Suppl. material 9: Table S6).

Several possible explanations exist for their non-detection/absence: first, the lack of detection could be attributed to the failure of the BF2/BR2 primer pair to effectively amplify some taxonomic groups, as has been found in previous studies (Creer et al. 2010; Elbrecht and Leese 2017), which also seems the most plausible explanation. Second, these taxa might have been absent from the sampling sites at the time of sampling for the metabarcoding study. However, although we cannot be fully sure that all three groups were present and sampled during the DNA metabarcoding fieldwork, the preliminary morphological sampling strongly supports this assumption. Third, there is a lack of public COI reference data for these taxa, hampering molecular species identification by automated comparison with available databases. However, if taxa were amplified, identification should have been possible at least to a higher taxonomic level, e.g. order. As this was not the case, it seems likely that these taxa were not amplified at all.

In addition to the detection of meiofaunal organisms, many macrozoobenthic organisms and several taxa not present in the preliminary morphological dataset were detected in the metabarcoding dataset. The detection of some taxa can likely be attributed to free DNA, cells or small body fragments of these taxa being present in the sampled sediments, e.g. *Homo sapiens*, *Gasterosteus aculeatus*, *Agelastica alni*, *Elodes minuta*, *Pediobius saulius*
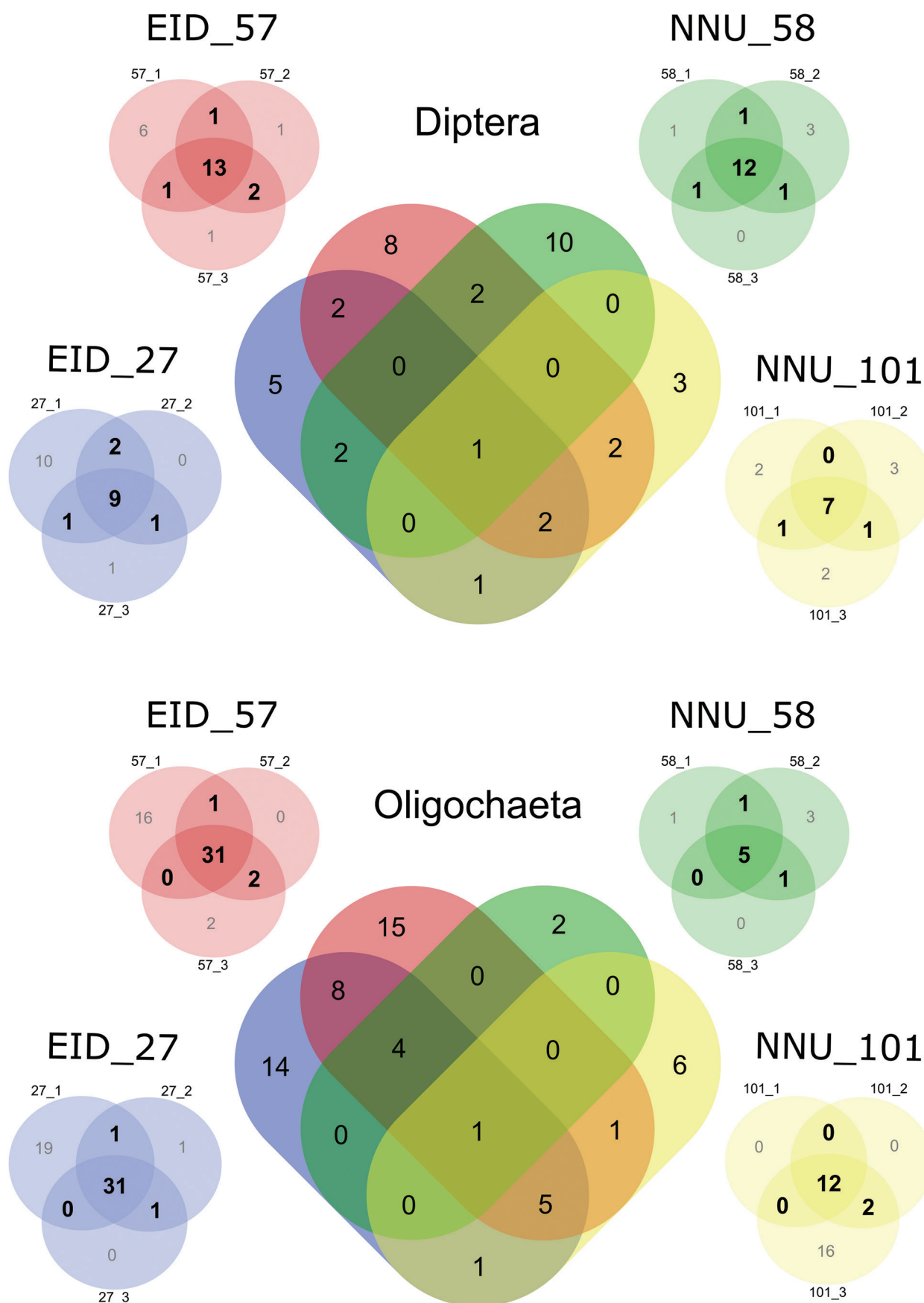
Figure 4. Overlap of MOTUs between PCR replicates and between sites for Diptera and Oligochaeta. Venn diagrams of Diptera and Oligochaeta MOTUs found per replicate, number of MOTUs shared between PCR replicates and number of MOTUs shared between sampling sites when only MOTUs present in at least two out of three replicates per sampling site were considered.

and Psocoptera. This corresponds to findings in previous studies showing that cells or DNA of metazoans can be extracted and amplified from soil samples (Bienert et al. 2012, Epp et al. 2012). Other taxa might have been present in the sediment as eggs. Prominent groups only identified by DNA metabarcoding were Cnidaria, Trichoptera, Plecoptera and Isopoda.

The most apparent advantage of the developed DNA metabarcoding approach for hyporheic communities is the detection of a large number of species that might not be found when applying standard kick-net sampling and omitting size sorting of samples to prevent over-representation of taxa with a high biomass per specimen (Elbrecht et al. 2017). However, only a fraction (17.5%) of all detected MOTUs could be assigned to a taxonomic name and only a very small fraction (1.2%) of the detected MOTUs could be reliably identified as Metazoa. When analysing the read abundance per taxonomic group, these differences were less pronounced, but still evident, as only 33.6% of all quality filtered reads belonged to the identified metazoan taxa. This ratio might be increased by applying flotation techniques that select for meiofaunal organisms, thus removing non-target taxa such as bacteria attached to sediment from the sample (Andújar et al. 2018, Burgess 2001, Haenel et al. 2017). The phenomenon that highly degenerate COI primers amplify a multitude of potential non-target taxa has been described before (Macher and Leese 2017, Siddall et al. 2009, Wangensteen et al. 2018) and designing primers targeting specific taxa of interest might be beneficial for future studies. Yet, the majority of studies applying the BF2/BR2 primer combination investigated (size-sorted) freshwater invertebrate bulk samples. In those studies, the proportion of target taxa (i.e. macroinvertebrates) and non-target organisms (e.g. bacteria from the cuticula, in gut contents or attached to the sediment) is much more beneficial for obtaining a high proportion of target MOTUs and sequence reads, compared to the size-sorted meiofaunal extraction approach applied here, focusing on the size fraction of <1 mm and >42 μm. Designing new primers targeting meiofaunal taxa commonly found in the hyporheic zone is promising, as currently available, highly degenerate COI primers have been designed for macroinvertebrate taxa (BF2/BR2: Elbrecht and Leese 2017) or marine invertebrates (Leray et al. 2013, Wangensteen et al. 2018). Therefore, they are more likely to miss taxa and lead to an underestimation of meiofaunal diversity, as seems partly the case in our study.

Further, the use of degenerate primers can help to detect and identify metazoan and non-metazoan MOTUs potentially suitable for biomonitoring and ecological studies. However, COI might not be the ideal marker gene for studying some metazoan (Creer et al. 2010) and especially non-metazoan taxa, as questions remain regarding its diagnostic power for species-level identification (Pawlowski et al. 2012). In addition, reference libraries for non-metazoan taxa contain mostly other genes, e.g. 18S or ITS (Kõljalg et al. 2005, Quast et al. 2012). Therefore,

further research is needed in order to assess the usability of COI-based MOTUs for ecosystem assessment. Machine-learning approaches as developed by Cordier et al. (2017) seem promising as they allow the use of MOTUs with unknown taxonomy and unknown ecological traits for ecosystem assessment. Yet, it must be highlighted that (Linnaean) taxonomy-free MOTU-based diversity analyses are influenced by the underlying data quality filtering settings, MOTU clustering approaches and post-clustering algorithms (e.g. see Coissac et al. 2012, Boyer et al. 2014, Mahé et al. 2015, Frøslev et al. 2017).

Site heterogeneity and metazoan community composition

Beta diversity and, as such, site heterogeneity were considerably high for Metazoa, but also when investigating Diptera and Oligochaeta only. In total, 70% of all identified metazoan taxa (126/180) were exclusively present at a single sampling site each, despite pooling nine sediment cores at each sampling site. Thereby, the finding that beta diversity estimates between sites increased when only MOTUs found in at least two out of three replicates were included in analyses could hint to a real biological phenomenon: taxa abundant at one sampling site might be rare at another site, therefore being removed by restrictive MOTU filtering. Another explanation would be the presence of taxa due to contamination or tag switching, which is a common phenomenon on Illumina sequencing platforms (Schnell et al. 2015). In the latter case, removing rare MOTUs would increase the reliability of study results, as the retained community would more accurately depict the real community present at a sampling site.

The findings that EID sites harboured more taxa than NNU sites and that communities in EID sites were more similar than communities in NNU sites could be explained by different phenomena. It might show that upstream sites are more isolated from each other, therefore harbouring more distinct metazoan communities. In addition, very small streams in upstream sites tend to dry out during times of low rainfall, therefore communities in these streams might be adapted to more extreme conditions and might more frequently face community fluctuations. In contrast, communities in downstream sites are expected to be more often connected, receive drifting organisms from different upstream regions and rarely become dry and hence, communities could be more constant and similar in EID sites. The finding that freshwater communities are more different in upstream and generally isolated sites has been reported before (Altermatt et al. 2013, Finn and Poff 2005). However, the small number of sites sampled in this study does not allow answering the question if the same patterns are consistently found when studying hyporheic communities and further research is needed. Nevertheless, it seems obvious that, in particular, oligochaetes and dipterans drive the observed patterns in our study.

The impact of replication strategy on inferred community composition

The results of our study highlight that replication strategy and removal of MOTUs present in only one PCR replicate can greatly impact the inferred diversity within and across sampling sites. Although the dataset analysed in this study is small and does not allow drawing definite conclusions, the results suggest that replication and removal of rare MOTUs found only in one out of three replicates could be beneficial for future studies and monitoring of hyporheic communities. As previously outlined in Alberdi et al. (2017), this approach is a compromise between retaining potential false-positives and introducing potential false-negatives. Analyses of community similarity between replicates showed that the removal of rare MOTUs leads to greatly increased similarity between replicates and therefore, false-positives are expected to be efficiently removed from the dataset. However, it is possible that the number of false-negatives increases and that some false-positives with a high read abundance remain in the dataset. Our results further suggest that for studies focusing on comparing (hyporheic meiofaunal) biodiversity in larger areas and across multiple sampling sites, PCR replication might not be that important, as beta diversity estimates remained largely independent of replication and removal of rare MOTUs. However, dissimilarity between sampling sites increased when only MOTUs present in at least two out of three replicates were considered. Therefore, in future studies, the best replication and filtering approach should be chosen based on the research question/application scenario, available financial resources and sequencing depth.

Future application scenarios

Hyporheic communities of stream ecosystems are often highly diverse. At the same time, taxonomic keys for genus and/or species level identification are widely lacking (Robertson et al. 2000). DNA-based approaches such as DNA metabarcoding have the potential for circumventing this shortcoming (Creer et al. 2010) by referring to the digitally preserved taxonomic knowledge stored in DNA barcode reference libraries. Given the steady progress made in the setup of those libraries (Ratnasingham and Hebert 2007), it can be expected that many taxa will be reliably identifiable on species level in the future – in case libraries are well curated. Even if species level identification is not possible, identification of MOTUs can be used to infer the ecology of taxa (Macher et al. 2016) and therefore, metabarcoding approaches hold great promises for the assessment of ecosystems and study of taxa that are difficult to identify based on their morphology. Hence, metabarcoding of hyporheic freshwater communities has great potential for diverse application scenarios such as the following:

Stream restoration management: The hyporheic zone is often rapidly colonised by meiofaunal organisms (Eisendle-Flöckner et al. 2013). Therefore, using metabarcoding to characterise the hyporheic community can potentially provide information on the ecological status of restored streams soon after restoration and before the arrival of classical bioindicator taxa such as Ephemeroptera, Trichoptera and Plecoptera (EPT taxa).

Streambed health assessment: Studies have shown that substrate characteristics are highly important for ecological processes in streams and freshwater communities are strongly influenced by substrate characteristics (Duan et al. 2008, Radwell and Brown 2006, 2007). When the ecological demands of taxa are known, metabarcoding hyporheic communities can potentially provide information on the ecological quality and function of the streambed. This might be particularly attractive in the context of monitoring streambed clogging, which can reduce accessibility and oxygenation of the hyporheic zone and subsequently leads to the loss of fish spawning grounds and habitats for species such as lampreys (Kemp et al. 2011). In particular, a high diversity of oligochaetes, as observed in the current study, might be an indication for intense streambed clogging (Bo et al. 2007).

Improvement of existing indices: As hyporheic fauna composition largely depends on small-scale environmental factors such as streambed substratum, a high resolution for ecological health assessments of streams could be reached. Some existing indices use information on certain taxa to assess ecosystem quality, e.g. the Oligochaete Index of Sediment Bioindication (IOBS; Rosso et al. 1994) and the Nematode Species at Risk Index (NemaSPEAR; Höss et al. 2011). Metabarcoding of hyporheic freshwater communities using taxon-specific primers could help to increase the accuracy of such indices by allowing for species level identification of otherwise difficult to identify taxa, as proposed by, for example, Vivien et al. (2015, 2017).

Assessment of intermittent stream and spring ecosystems: Standard assessment protocols for streams require the presence of water. However, intermittent rivers are common in many arid parts of the world and the hyporheic zone is an important refuge for aquatic taxa during droughts (Robson et al. 2011, Vander Vorste et al. 2016). Metabarcoding the hyporheic community might therefore help monitoring and assessing the ecological status of intermittent rivers throughout the year (Stubbington et al. 2018). The same applies for springs (i.e. ecotones between surface waters and groundwater aquifers), which can move in their longitudinal position over the year and depending on aquifer discharge conditions.

All those aspects are currently tackled in the two EU COST Actions 'DNAqua-Net - Developing new genetic tools for bioassessment of aquatic ecosystems in Europe' (Leese et al. 2016, 2018) and 'SMIRES - Science and Management of Intermittent Rivers and Ephemeral Streams' (Datry et al. 2017).

Conclusions

Highly degenerate COI primers are known to amplify a multitude of non-metazoan organisms, especially

when applied to samples with a low ratio of metazoan to non-metazoan taxa (Macher and Leese 2017, Siddall et al. 2009, Wangensteen et al. 2018). Our study indicates that this is also the case for hyporheic communities when using the BF2/BR2 primer pair for the investigation of stream meiofauna. Therefore, in future studies, appropriate primers should be chosen based on the expected and targeted taxonomic groups and depending on the research questions/application scenarios at hand. While the utilisation of highly degenerate primers has the benefit of amplifying a wide range of taxonomic groups, the disadvantage is that many reads can be "lost" to non-target, non-metazoan taxa that might be out of focus (such as bacteria). However, for widely understudied ecosystems like the hyporheic zone, using highly degenerate primers allows first insights into species diversity and biomonitoring can potentially benefit from including so far understudied and neglected non-metazoan taxa. Based on the initial results obtained by metabarcoding with highly degenerate primers, less degenerate primers can be chosen or designed for future studies, targeting taxonomic groups of interest.

Data resources

Raw data is available from the Short Read Archive (SRA), accession number SRR7623732.

Author contributions

AMW and JNM designed the study, performed sampling, analysed the data and wrote the manuscript.

Acknowledgements

References

Alberdi A, Aizpurua O, Gilbert MTP, Bohmann K (2017) Scrutinizing key steps for reliable metabarcoding of environmental samples. Methods in Ecology and Evolution 9: 134–147. https://doi.org/10.1111/2041-210X.12849

Altermatt F, Seymour M, Martinez N (2013) River network properties shape α-diversity and community similarity patterns of aquatic insect communities across major drainage basins. Journal of Biogeography 40: 2249–2260. https://doi.org/10.1111/jbi.12178

Andújar C, Arribas P, Gray C, Bruce C, Woodward G, Yu DW, Vogler AP (2018) Metabarcoding of freshwater invertebrates to detect the effects of a pesticide spill. Molecular Ecology 27: 146–166. https://doi.org/10.1111/mec.14410

Arribas P, Andújar C, Hopkins K, Shepherd M, Vogler AP (2016) Metabarcoding and mitochondrial metagenomics of endogean arthropods to unveil the mesofauna of the soil. Methods in Ecology and Evolution 7: 1071–1081. https://doi.org/10.1111/2041-210X.12557

Avó AP, Daniell TJ, Neilson R, Oliveira S, Branco J, Adão H (2017) DNA barcoding and morphological identification of benthic nematodes assemblages of estuarine intertidal sediments: advances in molecular tools for biodiversity assessment. Frontiers in Marine Science, 4, 66. https://doi.org/10.3389/fmars.2017.00066

Beier S, Traunspurger W (2001) The meiofauna community of two small German streams as indicator of pollution. Journal of Aquatic Ecosystem Stress and Recovery 8: 387–405. https://doi.org/10.1023/A:1012965424527

Benson DA, Cavanaugh M, Clark K, Karsch-Mizrachi I, Ostell J, Pruitt KD, Sayers EW (2017) GenBank. Nucleic Acids Research 46: D41–D47. https://doi.org/10.1093/nar/gkw1070

Bienert F, De Danieli S, Miquel C, Coissac E, Poillot C, Brun J-J, Taberlet P (2012) Tracking earthworm communities from soil DNA. Molecular Ecology 21: 2017–2030. https://doi.org/10.1111/j.1365-294X.2011.05407.x

Birk S, Bonne W, Borja A, Brucet S, Courrat A, Poikane S, Hering D (2012) Three hundred ways to assess Europe's surface waters: An almost complete overview of biological methods to implement the Water Framework Directive. Ecological Indicators 18: 31–41. https://doi.org/10.1016/j.ecolind.2011.10.009

Bista I, Carvalho GR, Walsh K, Seymour M, Hajibabaei M, Lallias D, Creer S (2017) Annual time-series analysis of aqueous eDNA reveals ecologically relevant dynamics of lake ecosystem biodiversity. Nature communications 8: 14087. https://doi.org/10.1038/ncomms14087

Blackman R, Constable D, Hahn C, Sheard A, Durkota J, Hänfling B, Handley LL (2017) Detection of a new non-native freshwater species by DNA metabarcoding of environmental samples – first record of *Gammarus fossarum* in the UK. Aquatic Invasions 12: 177–189. https://doi.org/10.3391/ai.2017.12.2.06

Bo T, Fenoglio S, Malacarne G, Pessino M, Sgariboldi F (2007) Effects of clogging on stream macroinvertebrates: an experimental approach. Limnologica-Ecology and Management of Inland Waters 37: 186–192. https://doi.org/10.1016/j.limno.2007.01.002

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: 176–182. https://doi.org/10.1111/1755-0998.12428

Burgess R (2001) An improved protocol for separating meiofauna from sediments using colloidal silica sols. Marine Ecology Progress Series 214: 161–165. https://doi.org/10.3354/meps214161

Coissac E, Riaz T, Puillandre N (2012) Bioinformatic challenges for DNA metabarcoding of plants and animals. Molecular Ecology 21: 1834–1847. https://doi.org/10.1111/j.1365-294X.2012.05550.x

Cordier T, Esling P, Lejzerowicz F, Visco J, Ouadahi A, Martins C, Pawlowski J (2017) Predicting the Ecological Quality Status of Marine Environments from eDNA Metabarcoding Data Using Supervised Machine Learning. Environmental Science & Technology 51: 9118–9126. https://doi.org/10.1021/acs.est.7b01518

Creer S, Fonseca VG, Porazinska DL, Giblin-Davis RM, Sung W, Power DM, Thomas WK (2010) Ultrasequencing of the meiofaunal biosphere: practice, pitfalls and promises. Molecular Ecology 19: 4–20. https://doi.org/10.1111/j.1365-294X.2009.04473.x

Datry T, Bonada N, Boulton AJ (2017) Intermittent Rivers and Ephemeral Streams: Ecology and Management. Hrvatske vode 25: 102.

Datry T, Singer G, Sauquet E, Jorda-Capdevila D, Von Schiller D, Subbington R, Zoppini A (2017) Science and Management of Intermittent Rivers and Ephemeral Streams (SMIRES). Research Ideas and Outcomes 3: e21774. https://doi.org/10.3897/rio.3.e21774

Duan X, Wang Z, Tian S (2008) Effect of streambed substrate on macroinvertebrate biodiversity. Frontiers of Environmental Science & Engineering in China 2: 122–128. https://doi.org/10.1007/s11783-008-0023-y

Edgar RC (2010) Search and clustering orders of magnitude faster than BLAST. Bioinformatics 26: 2460–2461. https://doi.org/10.1093/bioinformatics/btq461

Edgar RC (2013) UPARSE: highly accurate OTU sequences from microbial amplicon reads. Nature Methods 10: 996–998. https://doi.org/10.1038/nmeth.2604

Eisendle-Flöckner U, Jersabek CD, Kirchmair M, Hashold K, Traunspurger W (2013) Community patterns of the small riverine benthos within and between two contrasting glacier catchments. Ecology and Evolution 3: 2832–2844. https://doi.org/10.1002/ece3.679

Elbrecht V, Leese F (2015) Can DNA-Based Ecosystem Assessments Quantify Species Abundance? Testing Primer Bias and Biomass-Sequence Relationships with an Innovative Metabarcoding Protocol. PLoS One 10: e0130324. https://doi.org/10.1371/journal.pone.0130324

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

Elbrecht V, Peinert B, Leese F (2017) Sorting things out: Assessing effects of unequal specimen biomass on DNA metabarcoding. Ecology and Evolution 7: 6918–6926. https://doi.org/10.1002/ece3.3192

Elbrecht V, Taberlet P, Dejean T, Valentini A, Usseglio-Polatera P, Beisel J-N, Leese F (2016) Testing the potential of a ribosomal 16S marker for DNA metabarcoding of insects. PeerJ 4: e1966. https://doi.org/10.7717/peerj.1966

Elbrecht V, Vamos E, Meissner K, Aroviita J, Leese F (2017) Assessing strengths and weaknesses of DNA metabarcoding based macroinvertebrate identification for routine stream monitoring. Methods in Ecology and Evolution 8: 1265–1275. https://doi.org/10.1111/2041-210X.12789

Epp LS, Boessenkool S, Bellemain EP, 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: 1821–1833. https://doi.org/10.1111/j.1365-294X.2012.05537.x

Ficetola GF, Pansu J, Bonin A, Coissac E, Giguet-Covex C, De Barba M, Taberlet P (2015) Replication levels, false presences and the estimation of the presence/absence from eDNA metabarcoding data. Molecular Ecology Resources 15: 543–556. https://doi.org/10.1111/1755-0998.12338

Finn DS, Leroy Poff N (2005) Variability and convergence in benthic communities along the longitudinal gradients of four physically similar Rocky Mountain streams. Freshwater Biology 50: 243–261. https://doi.org/10.1111/j.1365-2427.2004.01320.x

Fonseca VG, Carvalho GR, Nichols B, Quince C, Johnson HF, Neill SP, Creer S (2014) Metagenetic analysis of patterns of distribution and diversity of marine meiobenthic eukaryotes. Global Ecology and Biogeography: A Journal of Macroecology 23: 1293–1302. https://doi.org/10.1111/geb.12223

Fraser BG, Williams DD (1997) Accuracy and precision in sampling hyporheic fauna. Canadian Journal of Fisheries and Aquatic Sciences. Journal Canadien Des Sciences Halieutiques et Aquatiques 54: 1135–1141. https://doi.org/10.1139/f97-024

Frie C (2016) Meiofaunal Diversity in ecologically improved river sections under the effect of connectedness to near-natural headwater sections (B.Sc. thesis). Ruhr-University Bochum (Germany), 54 pp.

Frøslev TG, Kjøller R, Bruun HH, Ejrnæs R, Brunbjerg AK, Pietroni C, Hansen AJ (2017) Algorithm for post-clustering curation of DNA amplicon data yields reliable biodiversity estimates. Nature Communications 8: 1188. https://doi.org/10.1038/s41467-017-01312-x

Gibson JF, Shokralla S, Curry C, Baird DJ, Monk WA, King I, Hajibabaei M (2015) Large-Scale Biomonitoring of Remote and Threatened Ecosystems via High-Throughput Sequencing. PLoS One 10: e0138432. https://doi.org/10.1371/journal.pone.0138432

Giere O (2008). Meiobenthology: The Microscopic Motile Fauna of Aquatic Sediments. Springer Science & Business Media.

Guardiola M, Uriz MJ, Taberlet P, Coissac E, Wangensteen OS, Turon X (2015) Deep-Sea, Deep-Sequencing: Metabarcoding Extracellular DNA from Sediments of Marine Canyons. PLoS One 10: e0139633. https://doi.org/10.1371/journal.pone.0139633

Haase P, Pauls SU, Schindehütte K, Sundermann A (2010) First audit of macroinvertebrate samples from an EU Water Framework Directive monitoring program: human error greatly lowers precision of assessment results. Journal of the North American Benthological Society 29: 1279–1291. https://doi.org/10.1899/09-183.1

Haenel Q, Holovachov O, Jondelius U, Sundberg P, Bourlat SJ (2017) NGS-based biodiversity and community structure analysis of meiofaunal eukaryotes in shell sand from Hållö island, Smögen, and soft mud from Gullmarn Fjord, Sweden. Biodiversity Data Journal 5: e12731. https://doi.org/10.3897/BDJ.5.e12731

Hakenkamp CC, Palmer MA (2000) The Ecology of Hyporheic Meiofauna. In Streams and Ground Waters 307–336. https://doi.org/10.1016/B978-012389845-6/50014-4

Hebert PDN, Ratnasingham S, deWaard JR (2003) Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. Proceedings of the Royal Society of London B: Biological Sciences 270: 96–S99. https://doi.org/10.1098/rsbl.2003.0025

Höss S, Claus E, Von der Ohe PC, Brinke M, Güde H, Heininger P, Traunspurger W (2011) Nematode species at risk – A metric to assess pollution in soft sediments of freshwaters. Environment International 37: 940–949. https://doi.org/10.1016/j.envint.2011.03.013

Huson DH, Auch AF, Qi J, Schuster SC (2007) MEGAN analysis of metagenomic data. Genome Research 17: 377–386. https://doi.org/10.1101/gr.5969107

Kemp P, Sear D, Collins A, Naden P, Jones I (2011) The impacts of fine sediment on riverine fish. Hydrological Processes 25: 1800–1821. https://doi.org/10.1002/hyp.7940

Kõljalg U, Larsson K-H, Abarenkov K, Henrik Nilsson R, Alexander IJ, Eberhardt U, Vrålstad T (2005) UNITE: a database providing web-based methods for the molecular identification of ectomycorrhizal fungi. The New Phytologist 166: 1063–1068. https://doi.org/10.1111/j.1469-8137.2005.01376.x

Lawrence JE, Skold ME, Hussain FA, Silverman DR, Resh VH, Sedlak DL, McCray JE et. al (2013) Hyporheic Zone in Urban Streams: A Review and Opportunities for Enhancing Water Quality and Improving Aquatic Habitat by Active Management. Environmental Engineering Science 30: 480–501. https://doi.org/10.1089/ees.2012.0235

Leese F, Altermatt F, Bouchez A, Ekrem T, Hering D, Meissner K, ... & Zimmermann J (2016) DNAqua-Net: Developing new genetic tools for bioassessment and monitoring of aquatic ecosystems in Europe. Research Ideas and Outcomes 2: e11321. https://doi.org/10.3897/rio.2.e11321

Leese F, Bouchez A, Abarenkov K, Altermatt F, Borja Á, Bruce K, ... & Weigand AM (2018) Why We Need Sustainable Networks Bridging Countries, Disciplines, Cultures and Generations for Aquatic Biomonitoring 2.0: A Perspective Derived From the DNAqua-Net COST Action. Advances in Ecological Research 58: 63–99.

Leray M, Yang JY, Meyer CP, Mills SC, Agudelo N, Ranwez V, Machida RJ (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: 34. https://doi.org/10.1186/1742-9994-10-34

Macher J-N, Leese F (2017) Environmental DNA metabarcoding of rivers: Not all eDNA is everywhere, and not all the time. bioRxiv https://doi.org/10.1101/164046

Macher J-N, Salis RK, Blakemore KS, Tollrian R, Matthaei CD, Leese F (2016) Multiple-stressor effects on stream invertebrates: DNA barcoding reveals contrasting responses of cryptic mayfly species. Ecological Indicators 61: 159–169. https://doi.org/10.1016/j.ecolind.2015.08.024

Mahé F, Rognes T, Quince C, de Vargas C, Dunthorn M (2015) Swarm v2: highly-scalable and high-resolution amplicon clustering. PeerJ 3: e1420. https://doi.org/10.7717/peerj.1420

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

Oksanen J, Kindt R, Legendre P, O'Hara B, Stevens MHH, Oksanen MJ, Suggests MASS (2007) The vegan package. http://ftp.uni-bayreuth.de/math/statlib/R/CRAN/doc/packages/vegan.pdf

Pacioglu O (2009) Ecology of the hyporheic zone: a review. Cave and Karst Science 36: 69–76.

Pawlowski J, Audic S, Adl S, Bass D, Belbahri L, Berney C, de Vargas C (2012) CBOL Protist Working Group: Barcoding Eukaryotic Richness beyond the Animal, Plant, and Fungal Kingdoms. PLoS Biology 10(11): e1001419. https://doi.org/10.1371/journal.pbio.1001419

Perini K (2016) Emscher River, Germany - Strategies and Techniques. In Urban Sustainability and River Restoration 151–159. https://doi.org/10.1002/9781119245025.ch14

Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Glöckner FO (2012) The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Research 41: D590–D596. https://doi.org/10.1093/nar/gks1219

Radwell AJ, Brown AV (2006) Influence of fine sediments on meiofauna colonization densities in artificial stream channels. Archiv für Hydrobiologie 165: 63–75. https://doi.org/10.1127/0003-9136/2006/0165-0063

Radwell AJ, Brown AV (2007) Benthic meiofauna assemblage structure of headwater streams: density and distribution of taxa relative to substrate size. Aquatic Ecology 42: 405–414. https://doi.org/10.1007/s10452-007-9108-0

Ratnasingham S, Hebert PDN (2007) BOLD: The Barcode of Life Data System (http://www.barcodinglife.org). Molecular Ecology Notes 7: 355–364. https://doi.org/10.1111/j.1471-8286.2007.01678.x

Robertson AL, Rundle SD, Schmid-Araya JM (2000) Putting the meio- into stream ecology: current findings and future directions for lotic meiofaunal research. Freshwater Biology 44: 177–183. https://doi.org/10.1046/j.1365-2427.2000.00592.x

Robson BJ, Chester ET, Austin CM (2011) Why life history information matters: drought refuges and macroinvertebrate persistence in non-perennial streams subject to a drier climate. Marine and Freshwater Research 62: 801. https://doi.org/10.1071/MF10062

Rosso A, Lafont M, Exinger A (1994) Impact of Heavy Metals on Benthic Oligochaete Communities in the River Ill and Its Tributaries. Water Science and Technology: A Journal of the International Association on Water Pollution Research 29: 241–248. https://doi.org/10.2166/wst.1994.0111

Schnell IB, Bohmann K, Gilbert MTP (2015) Tag jumps illuminated - reducing sequence-to-sample misidentifications in metabarcoding studies. Molecular Ecology Resources 15: 1289–1303. https://doi.org/10.1111/1755-0998.12402

Semrau M, Hurck R (2012) Multi-Purpose, Interlinked and Without Barriers: The Emscher River Ecological Concept. River Conservation and Management: 243–247.

Siddall ME, Fontanella FM, Watson SC, Kvist S, Erséus C (2009) Barcoding bamboozled by bacteria: convergence to metazoan mitochondrial primer targets by marine microbes. Systematic Biology 58: 445–451. https://doi.org/10.1093/sysbio/syp033

Smith DP, Peay KG (2014) Sequence Depth, Not PCR Replication, Improves Ecological Inference from Next Generation DNA Sequencing. PLoS One 9: e90234. https://doi.org/10.1371/journal.pone.0090234

Stubbington R, Chadd R, Cid N, Csabai Z, Miliša M, Morais M, Datry T (2018) Biomonitoring of intermittent rivers and ephemeral streams in Europe: Current practice and priorities to enhance ecological status assessments. The Science of the Total Environment 618: 1096–1113. https://doi.org/10.1016/j.scitotenv.2017.09.137

Stubbington R, Dole-Olivier M-J, Galassi DMP, Hogan J-P, Wood PJ (2016) Characterization of Macroinvertebrate Communities in the Hyporheic Zone of River Ecosystems Reflects the Pump-Sampling Technique Used. PLoS One 11: e0164372. https://doi.org/10.1371/journal.pone.0164372

Sunnucks P, Hales DF (1996) Numerous transposed sequences of mitochondrial cytochrome oxidase I-II in aphids of the genus *Sitobion* (Hemiptera: Aphididae). Molecular Biology and Evolution 13: 510–524. https://doi.org/10.1093/oxfordjournals.molbev.a025612

Taberlet P, Coissac E, Pompanon F, Brochmann C, Willerslev E (2012) Towards next-generation biodiversity assessment using DNA metabarcoding. Molecular Ecology 21: 2045–2050. https://doi.org/10.1111/j.1365-294X.2012.05470.x

Tonscheidt A (2016) Meiofauna-Diversity in ecologically improved river sections - the effect of time (B.Sc. thesis). Ruhr-University Bochum (Germany), 44 pp.

Valentini A, Pompanon F, Taberlet P (2009) DNA barcoding for ecologists. Trends in Ecology & Evolution 24: 110–117. https://doi.org/10.1016/j.tree.2008.09.011

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

Vander Vorste R, Mermillod-Blondin F, Hervant F, Mons R, Datry T (2016) *Gammarus pulex* (Crustacea: Amphipoda) avoids increasing water temperature and intraspecific competition through vertical migration into the hyporheic zone: a mesocosm experiment. Aquatic Sciences 79: 45–55. https://doi.org/10.1007/s00027-016-0478-z

Vasselon V, Rimet F, Tapolczai K, Bouchez A (2017) Assessing ecological status with diatoms DNA metabarcoding: Scaling-up on a WFD monitoring network (Mayotte island, France). Ecological Indicators 82: 1–12. https://doi.org/10.1016/j.ecolind.2017.06.024

Vivien R, Holzmann M, Werner I, Pawlowski J, Lafont M, Ferrari BJD (2017) Cytochrome c oxidase barcodes for aquatic oligochaete identification: development of a Swiss reference database. PeerJ 5: e4122. https://doi.org/10.7717/peerj.4122

Vivien R, Wyler S, Lafont M, Pawlowski J (2015) Molecular barcoding of aquatic oligochaetes: implications for biomonitoring. PLoS One 10: e0125485. https://doi.org/10.1371/journal.pone.0125485

Wangensteen OS, Palacín C, Guardiola M, Turon X (2018) DNA metabarcoding of littoral hard-bottom communities: high diversity and database gaps revealed by two molecular markers. PeerJ 6: e4705. https://doi.org/10.7717/peerj.4705

Wangensteen OS, Turon X (2017) Metabarcoding techniques for assessing biodiversity of marine animal forests. In: Rossi S, Bramanti L, Gori A, Orejas C (Eds) Marine animal forests, the ecology of benthic biodiversity hotspots. Springer, 445–473. https://doi.org/10.1007/978-3-319-21012-4_53

Winking C, Lorenz AW, Sures B, Hering D (2014) Recolonisation patterns of benthic invertebrates: a field investigation of restored former sewage channels. Freshwater Biology 59: 1932–1944. https://doi.org/10.1111/fwb.12397

Yu DW, Ji Y, Emerson BC, Wang X, Ye C, Yang C, Ding Z (2012) Biodiversity soup: metabarcoding of arthropods for rapid biodiversity assessment and biomonitoring. Methods in Ecology and Evolution 3: 613–623. https://doi.org/10.1111/j.2041-210X.2012.00198.x

Supplementary material 1
DNA extraction protocol
Authors: Alexander M. Weigand, Jan-Niklas Macher
Link: https://doi.org/10.3897/mbmg.2.26869.suppl1

Supplementary material 2
Figure S1: Overview of sediment characteristics
Authors: Alexander M. Weigand, Jan-Niklas Macher
Link: https://doi.org/10.3897/mbmg.2.26869.suppl2

Supplementary material 3
Figure S2: Taxonomic composition of hyporheic community with all metazoan MOTUs
Authors: Alexander M. Weigand, Jan-Niklas Macher
Link: https://doi.org/10.3897/mbmg.@.26869.suppl3

Supplementary material 4
Table S1: Sampling site coordinates
Authors: Alexander M. Weigand, Jan-Niklas Macher
Link: https://doi.org/10.3897/mbmg.2.26869.suppl4

Supplementary material 5
Table S2: Primer combinations
Authors: Alexander M. Weigand, Jan-Niklas Macher
Link: https://doi.org/10.3897/mbmg.2.26869.suppl5

Supplementary material 6
Table S3: Overview of raw reads and quality filtered reads
Authors: Alexander M. Weigand, Jan-Niklas Macher
Link: https://doi.org/10.3897/mbmg.2.26869.suppl6

Supplementary material 7
Table S4: Metazoa MOTU list with all replicates and including site 61
Authors: Alexander M. Weigand, Jan-Niklas Macher
Link: https://doi.org/10.3897/mbmg.2.26869.suppl7

Supplementary material 8

Table S5: Metazoa MOTU list with 2-out-of-3 replicate strategy and excluding site 61

Authors: Alexander M. Weigand, Jan-Niklas Macher

Link: https://doi.org/10.3897/mbmg.2.26869.suppl8

Supplementary material 9

Table S6: Taxa identified based on morphology

Authors: Alexander M. Weigand, Jan-Niklas Macher

Link: https://doi.org/10.3897/mbmg.2.26869.suppl9