

Exploring the potential of metabarcoding to disentangle macroinvertebrate community dynamics in intermittent streams

Alex Laini¹, Arne J. Beermann^{2,3}, Rossano Bolpagni¹, Gemma Burgazzi¹, Vasco Elbrecht⁴, Vera M. A. Zizka⁴, Florian Leese^{2,3}, Pierluigi Viaroli¹

1 Department of Chemistry, Life Sciences and Environmental Sustainability, University of Parma, Parma, Italy

2 Aquatic Ecosystem Research, University of Duisburg-Essen, Essen, Germany

3 Centre for Water and Environmental Research (ZWU), University of Duisburg-Essen, Essen, Germany

4 Centre for Biodiversity Monitoring, Zoological Research Museum Alexander Koenig, Bonn, Germany

Corresponding author: Alex Laini (alex.laini@unipr.it)

Academic editor: Tristan Lefébure | Received 24 February 2020 | Accepted 5 August 2020 | Published 28 August 2020

Abstract

Taxonomic sufficiency represents the level of taxonomic detail needed to detect ecological patterns to a level that match the requirement of a study. Most bioassessments apply the taxonomic sufficiency concept and assign specimens to the family or genus level given time constraints and the difficulty to correctly identify species. This holds particularly true for stream invertebrates because small and morphologically similar larvae are hard to distinguish. Low taxonomic resolution may hinder detecting true community dynamics, which thus leads to incorrect inferences about community assembly processes. DNA metabarcoding is a new, affordable and cost-effective tool for the identification of multiple species from bulk samples of organisms. As it provides high taxonomic resolution, it can be used to compare results obtained from different identification levels. Measuring the effect of taxonomic resolution on the detection of community dynamics is especially interesting in extreme ecosystems like intermittent streams to test if species at intermittent sites are subsets of those from perennial sources or if independently recruiting taxa exist. Here we aimed to compare the performance of morphological identification and metabarcoding to detect macroinvertebrate community dynamics in the Trebbia River (Italy). Macroinvertebrates were collected from four perennial and two intermittent sites two months after flow resumption and before the next dry phase. The identification level ranged from family to haplotype. Metabarcoding and morphological identifications found similar alpha diversity patterns when looking at family and mixed taxonomic levels. Increasing taxonomic resolution with metabarcoding revealed a strong partitioning of beta diversity in nestedness and turnover components. At flow resumption, beta diversity at intermittent sites was dominated by nestedness when family-level information was employed, while turnover was evidenced as the most important component when using Operational Taxonomic Units (OTUs) or haplotypes. The increased taxonomic resolution with metabarcoding allowed us to detect species adapted to deal with intermittency, like the chironomid *Cricotopus bicinctus* and the ephemeropteran *Cloeon dipterum*. Our study thus shows that family and mixed taxonomic level are not sufficient to detect all aspects of macroinvertebrate community dynamics. High taxonomic resolution is especially important for intermittent streams where accurate information about species-specific habitat preference is needed to interpret diversity patterns induced by drying and the nestedness/turnover components of beta diversity are of interest to understand community assembly processes.

Key Words

beta diversity, biodiversity, exact sequence variants, IRES, nestedness, taxonomic sufficiency, turnover

Introduction

Community ecology aims to unravel principles that determine the generation, maintenance and distribution of diversity in space and time (Mittelbach and McGill

2019). This is particularly important in times of global change and biodiversity loss in order to put appropriate management measures in place (Wallington et al. 2005). Measuring community dynamics is especially interesting in extreme ecosystems, such as intermittent rivers and

ephemeral streams (IRES), to test if species at intermittent sites are subsets of those from nearby perennial sources (nestedness) or if they can autochthonously recruit organisms as the result of species-specific adaptations to lack of water (e.g. resistance forms on dry riverbeds; turnover) (Datry et al. 2017). IRES occur in any continent and are common in temperate regions and dominant in Mediterranean-climate and semi-arid river networks (Datry et al. 2017, Stubbington et al. 2018). The alternation of wet and dry periods in IRES shapes biotic communities by inducing multiple recolonisation events (Zickovich and Bohonak 2007), which strongly affect alpha and beta diversity. Under climate change, rivers and streams are predicted to become more intermittent. Therefore, understanding the drivers of community dynamics in IRES is a main challenge for freshwater science and a crucial one to support aquatic biodiversity management (Döll and Schmied 2012, Datry et al. 2016).

Inferences about community dynamics in IRES are usually based on data generated from morphological identification and frequently rely on taxonomic levels coarser than species as biodiversity surrogates. This practice has its roots in the concept of taxonomic sufficiency (TS), the taxonomic detail needed to meet the requirement of a study (Jones 2008, Vilmi et al. 2016). TS is a common practice in macroinvertebrate community research, even if highly diverse families contain species with different ecological requirements and high bioindicator value; for example, Chironomidae, Hydropsychidae and Baetidae (Resh and Unzicker 1975, Jones 2008, Lencioni et al. 2012). TS frequently represents the only option because diagnostic characters for morphological identification are complex, identification keys for species are lacking or damaged/immature specimens need to be identified. The use of TS is dictated by a trade-off between acquired information and feasibility in terms of money, time and lack of expertise (Jones 2008), despite species having to be considered as the gold standard for ecological studies (Gaston 2000b, Pawlowski et al. 2018). When using TS, one intrinsic assumption is that functional, biological and ecological traits are preserved at higher taxonomic levels. Studies on niche conservatism, however, show that closely related species can have very different niches (Knouft et al. 2006, Losos 2008). For instance, different species of the genus *Polypedilum* (Chironomidae) show contrasting responses to desiccation, with *P. vanderplanki* being desiccation-resistant via anhydrobiosis, while *P. nubifer* is desiccation-sensitive (Gusev et al. 2014). Similarly, Suemoto et al. (2004) demonstrate that various species of the genus *Chironomus* (Chironomidae) have evolved different strategies to resist desiccation.

In this context, new genetic techniques are emerging as affordable tools to overcome identification problems (Leese et al. 2016). DNA metabarcoding is a technique that allows species composition to be inferred from bulk samples (e.g. all the organisms found in a kick-net sample) by sequencing a specific gene region (Taberlet et al. 2012, Leese et al. 2016). Comparisons made between morpho-

logical identification and metabarcoding give comparable results for both ‘mock’ (communities of known composition) and ‘real’ communities, although some mismatches occur due to incomplete reference databases, primer bias and problems in detecting rare taxa (Hajibabaei et al. 2011, Elbrecht et al. 2017, Bush et al. 2019). However, metabarcoding typically achieves higher taxonomic resolution (Gibson et al. 2015, Elbrecht et al. 2017, Kuntke et al. 2020) and it allows to differentiate cryptic species that lack diagnostic characters but can respond differently to environmental stressors (Sturmbauer et al. 1999, Eisenring et al. 2016). Cryptic species, although lacking a formal description, can be assessed as Operational Taxonomic Units (OTUs) by relying on genetic data. OTUs can be used as a proxy of individual species and greatly enhance diagnostic power in environmental studies, as demonstrated by Beermann et al. (2018). Metabarcoding has been successfully used to detect the responses of macroinvertebrate communities to environmental gradients (Emilson et al. 2017) and pesticide spills (Andújar et al. 2018). Novel bioinformatic approaches also allow inferences of haplotypes from macroinvertebrate metabarcoding data (Elbrecht et al. 2018, Zizka et al. 2020). Although metabarcoding represents a cost- and time-efficient alternative to morphological identification (Elbrecht et al. 2017), its usefulness in studying macroinvertebrate dynamics in IRES needs to be tested (Stubbington et al. 2018). This is especially true when considering that metabarcoding does not provide abundance estimates (Elbrecht and Leese 2015) and may fail to recover small or rare taxa (Hajibabaei et al. 2011, Leray and Knowlton 2017).

The main aim of this work was, therefore, to evaluate the performances of different identification levels in detecting invertebrate community dynamics in IRES. More specifically, we compared the results obtained with metabarcoding and morphological identification and tested if an increased resolution from family to OTU and haplotype levels provides similar information about alpha diversity patterns, community structure and beta diversity partitioning in the turnover and nestedness components. We collected samples from the intermittent Trebbia River, a medium-sized tributary of the Po River (N Italy), 2 and 9 months after a dry event in 2017. These time steps were chosen because invertebrate community recovery to pre-drought conditions in intermittent streams with predictable seasonal dry periods ranges from weeks to few months (Lake 2003, Acuña et al. 2005, Datry 2012). We hypothesised that the macroinvertebrate community at intermittent sites represents a subset of the community present at perennial sites 2 months immediately after the dry period due to organisms drifting from upstream sites (‘nestedness’). We also hypothesised that the macroinvertebrate community completely recovers after 9 months from flow resumption with an increased turnover compared to perennial sites due to niche filtering, dispersal from other water bodies and competition amongst organisms. Metabarcoding and morphological identification are expected to provide similar patterns of alpha diversity,

community structure and turnover/nestedness dynamics, irrespective of the employed identification level (e.g. family, OTUs etc.). Identification of organisms at the species level is expected to provide information about the presence of specialist taxa that are able to survive drought with either resistance (e.g. desiccation resistance stages) or resilience (e.g. passive drift from upstream sites) strategies.

Material and methods

Study area

The Trebbia River is a 120 km-long tributary of the Po River with a mean annual discharge of nearly $20 \text{ m}^3 \text{ s}^{-1}$. Water is diverted for irrigation at $\sim 20 \text{ km}$ upstream from the confluence with the Po River and the environmental flow downstream of the last withdrawal is set at $1.5 \text{ m}^3 \text{ s}^{-1}$. In July and August, the median discharge upstream of the water withdrawal is $3.9 \text{ m}^3 \text{ s}^{-1}$, with an interquartile range of 2.3 and $5.7 \text{ m}^3 \text{ s}^{-1}$, respectively (data from the 2003–2015 period). The environmental flow does not generally suffice to feed the river up to the confluence with the Po River and the last $\sim 5 \text{ km}$ upstream of the confluence generally dry out in summer.

Six sampling sites from three different Trebbia River sections were selected (Fig. 1): i) upstream of the water abstraction - perennial (St1, St2); ii) downstream of the water abstraction - perennial (St3, St4); iii) downstream of the water abstraction - intermittent (St5, St6). Altitude ranges from 162 m a.s.l. (St1) to 49 m a.s.l. (St6). Intermittent sites experienced a dry period of nearly 2 months in summer 2017, during which the streambed completely dried. Macroinvertebrates were collected in December 2017, 2 months after flow resumption, and at the end of June/beginning of July (hereafter July) 2018, before the dry period. At each site, macroinvertebrates were collected at 10 locations in the main channel with a Surber net (0.05 m^2 frame area). Surber replicates were

pooled together and organisms were sorted directly in the field and preserved in absolute ethanol. Ethanol was replaced 1 day after collection and samples were preserved at $4 \text{ }^\circ\text{C}$. Organisms were morphologically identified at the lowest possible taxonomic level, usually genus or family, using a stereomicroscope according to Campaioli et al. (1994, 1999), Tachet (2010) and Miller and Bergsten (2016).

Metabarcoding

Metabarcoding was performed according to Elbrecht and Leese (2015). Briefly, each sample was dried overnight at room temperature and then ground with an IKA Ultra Turrax Tube Drive Control (Staufen, Germany) for 30 min at 4000 rpm. For each sample, only one part of the body of larger organisms (e.g. legs) was retained to increase the probability of detecting smaller organisms and rare species. DNA was extracted according to a modified salt extraction protocol (Weiss and Leese 2016). RNA was digested with RNase A to improve downstream analysis steps. DNA was purified using the NucleoSpin Gel and the PCR clean-up kit (Macherey and Nagel, Düren, Germany). DNA sample concentration was equilibrated to $25 \text{ ng } \mu\text{l}^{-1}$ and the first PCR step (20 cycles) was performed using the BF2+BR2 primers according to Elbrecht and Leese (2017) with the Qiagen Multiplex PCR Plus kit. Two replicates per sample were analysed to increase accuracy. Samples were multiplexed following the tagging strategy described in Elbrecht and Steinke (2019) in the second PCR step (15 cycles). The final library was prepared by pooling an equal amount of amplicons for each sample. The library was purified using $0.76\times$ SPRI select (Beckman Coulter, Krefeld, Germany) and sequenced with an Illumina HiSeq Rapid Run 2 \times 250 bp as part of a library containing nearly 280 samples at the Eurofins Genomics (Ebersberg, Germany). Illumina reads have been deposited in the SRA database under the accession number PRJNA655393.

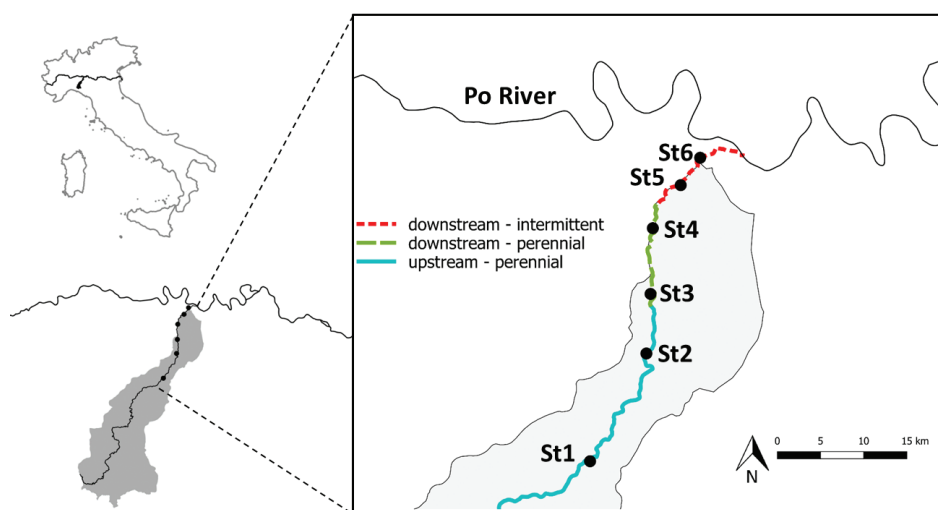


Figure 1. Map of the study area.

One of the most abundant species recovered by morphological identification, *Oligoneuriella rhenana*, was not detected by metabarcoding neither at species nor genus when a 97–95% cut-off was used for species/genus assignments. Single *O. rhenana* individuals were thus barcoded to check if any problems arose in the metabarcoding results (Suppl. material 2: Table S1, PCR settings upon request). All 24 *O. rhenana* barcode sequences were deposited in BOLD (sequence pages: TRBB001-20 - TRBB024-20).

The metabarcoding data were analysed with JAMP, version 0.69 (<https://github.com/VascoElbrecht/JAMP>), of the R software (R Core Team 2017). JAMP is an R-based metabarcoding pipeline that mostly relies on USEARCH 11.0.667 (Edgar 2010) and VSEARCH 2.10.4 (Rognes et al. 2016). First, the sequencing files were demultiplexed according to sample tags and quality-checked using FastQC. The paired-end reads were firstly merged and reverse complements were built whenever needed. Primers were trimmed with Cutadapt (Martin 2011) and sequences were filtered using a maximum expected error (max ee) filtering of 0.5 (Edgar and Flyvbjerg 2015), after which only those ± 10 bp of the expected length (421 bp) were retained. Given differences in sequencing depth, the number of reads of each sample was subset to the lowest number of reads found in our samples (547,000, site 1, 2nd replicate). OTU clustering was done by USEARCH with a clustering threshold of 97% similarity and reads, including singletons, matching the OTUs to generate an OTU table. The OTUs with a read abundance below 0.01% per sample were discarded, while only the OTUs present in both replicates were retained for the downstream data analysis. Finally, a taxonomic assignment was performed using BOLD (Ratnasingham and Hebert 2007), accessed on 07-05-2019. Besides OTUs, haplotypes were inferred by following the denoising approach as implemented in Elbrecht et al. (2018) using the default settings.

Data analysis

Pairwise correlations (Spearman's rank) were calculated to test if alpha diversity patterns were consistent amongst the identification levels (family, mixed, OTU and haplotype) and methods (morphology, metabarcoding). Organisms were identified to a mixed level for both identification methods and resulted in taxa being present at multiple levels of the taxonomic hierarchy. In our work, mixed identification levels were present because of immature/damaged specimens for morphological identification and because related sequences were assigned to different taxa with metabarcoding. This inconsistency can affect the overall result of the study because of ambiguous parent-child pairs being present (e.g. *Baetis* and Baetidae). We decided not to resolve ambiguous parent-child pairs because different methods can lead to different results (Cuffney et al. 2007). Taxonomic information was managed with the package 'biomonitoR' (<https://github.com/alexology/biomonitoR>).

Non-metric multidimensional scaling (nMDS) was performed with the metaMDS function of the R package 'vegan' (Oksanen et al. 2019) using different combinations of identification method and taxonomic level: i) mixed levels and family from the morphological identification using both abundance and presence/absence data; ii) mixed level and family from metabarcoding presence-absence data; iii) OTUs presence-absence data; iv) data on intraspecific diversity, i.e. haplotype presence-absence data. Jaccard and Bray-Curtis dissimilarities were used for presence-absence and abundance data, respectively. nMDS was followed by a PROTEST analysis which allows to test the non-randomness of the correlation between two ordinations that is bound between 0 (no fit) and 1 (perfect fit) (Peres-Neto and Jackson 2001). It allows to test if different taxonomic levels or identification methods provide similar community structure information. The turnover and nestedness components were then calculated to test if the estimated beta diversity components were affected by TS. Beta diversity partitioning was performed according to the approach proposed by Baselga (2010), where nestedness is calculated as the difference between total dissimilarity, estimated by the Sørensen index, and turnover, estimated by the Jaccard index. The contribution of turnover and nestedness was then calculated as percentages of total dissimilarity. The correlation of turnover percentage amongst taxonomic levels and identification methods was then assessed by a Mantel test based on Spearman's rank correlations. The Mantel tests obtained from nestedness and turnover contributions were equal because both were calculated as the percentage of the total dissimilarity. The probability values for alpha diversity, PROTEST and the Mantel test were corrected for multiple comparisons with the Holm correction. The effect of phylogenetic structure on the partitioning of beta-diversity was evaluated by partitioning the PhyloSor index (Bryant et al. 2008) in turnover and nestedness components according to Leprieur et al. (2012). At first, OTU and haplotype sequences were aligned using MUSCLE (Edgar 2004). Maximum likelihood trees were constructed from OTU and haplotype sequences using RAxML version 8 (Stamatakis 2014) called inside R with the 'ips' package (Heibl 2008). Briefly, we used the general time reversible model (GTRCAT) and used the automated, rapid bootstrapping analysis (autoMRE) to search for the bestscoring ML tree in one program run. Beta diversity partitioning analyses were performed with the package 'betapart' (Baselga et al. 2018). The package 'ggplot2' (Wickham 2009) was used to plot the results.

Results

Comparison between morphological and metabarcoding data

With the morphological approach 4050 organisms belonging to 69 taxa were identified, while 132 taxa, 229 OTUs and 513 haplotypes were detected by metabar-

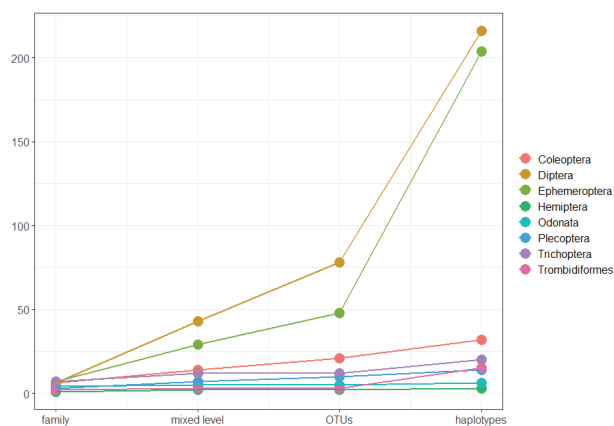


Figure 2. Richness from family to haplotypes for the main macroinvertebrate groups found in this study.

coding and bioinformatic protocols. According to morphological identification, the most abundant order was Ephemeroptera (1499 organisms), followed by Diptera (1157), Trichoptera (738), Trombidiformes (261), Plecoptera (165), Coleoptera (132), Odonata (56), Heteroptera (12) and other groups (30). Diptera and Ephemeroptera were the taxonomic groups that benefited most from the increase in the identification level in terms of richness (Fig. 2). Moreover, the abundance of the most abundant families of Diptera, Ephemeroptera, Trichoptera and Coleoptera was related to haplotype richness (Suppl. material 2: Fig. S1).

Fewer taxa than the number of OTUs were found because several OTUs were assigned to the same taxon, while others were not assigned at all. Table 1 reports the number of taxa found at the different taxonomic levels by both the morphological and metabarcoding approaches. Metabarcoding missed some taxa identified by the morphological approach at both the family and genus level. Families Blephariceridae, Ceratopogonidae, Dugesiidae, Philopotamidae, Physidae and Scirtidae were detected only by the morphological approach, while Hydracarina families Lebertiidae, Mideopsidae and Tetrastemmatidae were detected only by metabarcoding. The mean family richness obtained by the morphological approach (18.2 ± 7.0 , mean \pm standard deviation) was not significantly different (paired t-test, t-value = -1.63, df = 11, p-value = 0.13) from that inferred via metabarcoding (17.2 ± 5.8) (Fig. 3). The overlap between the morphological and metabarcoding results, calculated as the ratio between the number of shared families and the number of total families per site, was 84.2% on average, with a minimum of 68.4% and a maximum of 100%.

The genera *Esolus* (Coleoptera), *Wormaldia* (Trichoptera), *Dugesia* (Turbellaria) and *Dinocras* (Plecoptera) were not detected by metabarcoding. The sequences related to these genera were probably assigned to coarser taxonomic levels as our OTU table comprised unassigned Insecta, Coleoptera, Trichoptera, Turbellaria, Dryopidae and Perlidae (e.g. similarity to reference

Table 1. Richness obtained by morphological identification and metabarcoding at different identification levels. The metabarcoding data of this study were only analysed in terms of presence-absence.

Taxonomic level	Metabarcoding	Morphology
Phylum	4	4
Class	6	6
Order	11	8
Family	40	43
Genus	68	45
Species	82	0
Mixed taxonomic level	133	69
OTU	229	NA
Haplotype	513	NA

sequences > 10%). Metabarcoding originally identified the trichopteran species *Psychomyia pusilla* as *Rhyacophila pusilla*, which is only known from China (Malicky and Changhai 2002). This assignment proved to be a misclassification problem in the BOLD database and was corrected (records HMKKT101-10 and HM-KKT369-10 on the BOLD website). The only mollusc taxon collected in the Trebbia River (*Physa*) was not detected by metabarcoding, probably because of a primer bias (Elbrecht and Leese 2015) and Turbellaria were assigned only at the class level, probably because suitable reference data were lacking.

Barcoding of single *Oligoneuriella* specimens confirmed both the morphological and metabarcoding identifications. For Italy, the family Oligoneuriidae is represented by the species *Oligoneuriella rhenana*. The best match of our specimens to the *O. rhenana* sequences found in BOLD was 91.79% similarity.

Effects of identification method and taxonomic resolution on alpha diversity, community structure and beta diversity

A high rank correlation between the morphological- and metabarcoding-based alpha diversity estimates was found for both the mixed (0.88, $p < 0.01$) and family (0.93, $p < 0.001$) level. The high correlation for the mixed taxonomic level was present despite the different taxonomic levels used with the two identification methods. In metabarcoding, a high correlation was found for the mixed taxonomic level with OTUs (0.80, $p < 0.05$) and haplotypes (0.80, $p < 0.05$). Non-significant correlations were observed for the family level identification with OTUs (0.51, $p > 0.05$) and haplotypes (0.50, $p > 0.05$). All the results are reported in Table 2.

The community structure assessed with nMDS was similar at all the taxonomic levels investigated by the morphological and metabarcoding approaches (Fig. 4, Suppl. material 2: Fig. S2). The visual similarity between the nMDS ordinations was also supported by the PROTEST analyses, which revealed significant comparisons for all the analysed pairs (Table 3). However, while the correlation was greater than 0.88 for different Linnaean taxonomic levels (range 0.88–0.98), the correlation be-

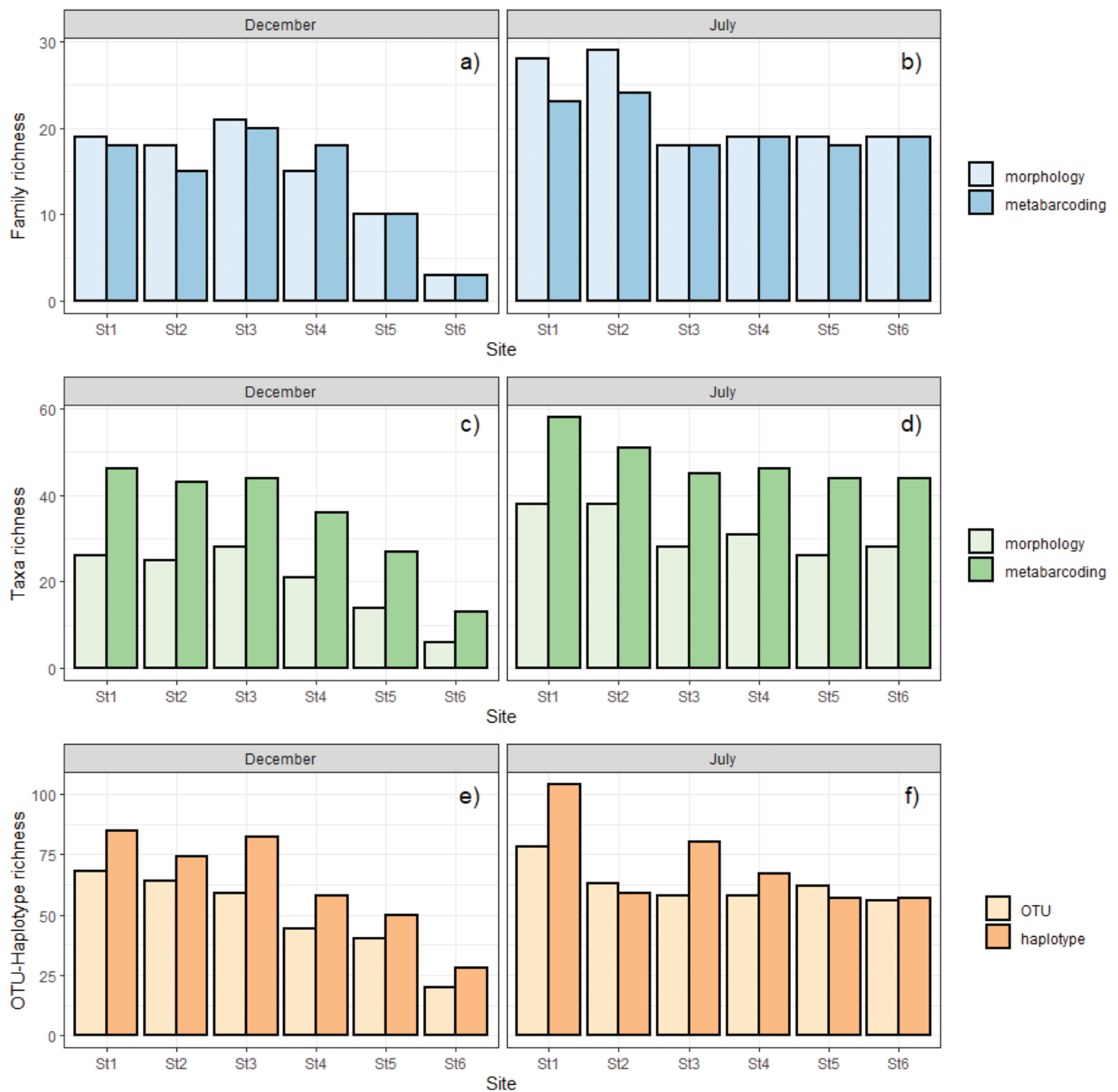


Figure 3. Family (a–b) and mixed taxonomic level (lowest possible level; c–d) richness for the morphological identification and metabarcoding in December 2017 (shortly after the dry season) and July 2018 (shortly before the next dry season). OTU and haplotype richness are also shown (e–f).

Table 2. Pairwise Spearman’s rank correlation coefficients of the alpha diversity performed at different identification levels (fam = family, taxa = mixed taxonomic level) for both methods (mor = morphological identification, meta = metabarcoding).

fam mor	0.93	0.88	0.86	0.72	0.54
***	fam meta	0.93	0.80	0.51	0.50
**	***	taxa mor	0.88	0.57	0.58
**	*	**	taxa meta	0.80	0.80
.			*	OTU	0.78
			*	*	haplotype

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, . $p < 0.1$

tween the OTUs/haplotypes and the Linnaean taxonomic levels dropped to range from 0.76 to 0.86. These findings indicate changes when using the highest, even intraspecific

(haplotype), resolution. The July and December samples formed two separate clusters in all nMDS analyses. Amongst the December samples, intermittent sites St5 and St6 deviated from the bulk of perennial sites. This pattern was consistent amongst all the tested taxonomic levels and also between these levels and the haplotype level, but was less distinct for the ordinations obtained with morphological data, where St5 was less separated from perennial sites (Figure 4a, b). Distance amongst the July samples on the ordination space was shorter than for the December ones and both perennial and intermittent sites tended to cluster together. Once again, the nMDS performed on OTU/haplotype data provided a slightly different interpretation, because intermittent St6 was very much isolated from the

Table 3. Correlation amongst the nMDS ordinations performed at different identification levels (fam = family, abu = abundance, pa = presence-absence, taxa = mixed taxonomic level) for both methods (mor = morphological identification, meta = metabarcoding).

fam abu mor	0.97	0.96	0.96	0.94	0.88	0.77	0.79
**	fam pa mor	0.98	0.96	0.95	0.9	0.81	0.81
**	**	fam meta	0.92	0.93	0.89	0.82	0.79
**	**	**	taxa abu mor	0.98	0.95	0.79	0.82
**	**	**	**	taxa pa mor	0.94	0.80	0.86
**	**	**	**	**	taxa meta	0.76	0.78
**	**	**	**	**	**	OTU	0.96
**	**	**	**	**	**	**	haplotype

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, . $p < 0.1$

**Figure 4.** nMDS ordination plots at different identification levels: family level for morphology (a) and metabarcoding (b) data, OTUs (c) and haplotypes (d). Stress was lower than 10% for all analyses. Numbers refer to sites: 1–2 upstream perennial, 3–4 downstream perennial and 5–6 downstream intermittent.

other sites (Figure 4c, d). The *metaMDS* function used to perform nMDS rotates the configuration to maximise the variance of points on the first dimension, so it was possible to directly compare the patterns highlighted by different ordinations. When we took the investigated sites as a continuum along the river, the nMDS performed on the Linnaean taxonomic levels showed a gradient orientated in a bottom-left to top-right direction. This was not the case for the nMDS performed on OTUs, where the gradient was orientated in a bottom-up direction. This scenario highlights a temporal trend on the first axis and a spatial trend on the second one. The nMDS performed on the abundance data behaved similarly to that performed with the presence-absence data. However, the nMDS performed at the family level with the morphological approach identified a subgroup formed by the perennial sites located upstream of the water abstraction (St1 and St2) in July, while the others did not.

Lowering the taxonomic level had a huge impact on beta diversity, which increased on average when taxonomic resolution rose (Fig. 5, Suppl. material 2: Fig. S3). The percentages of the nestedness and turnover components of beta diversity reversed by going from a higher (family) to a lower (haplotype) identification level, at least for the December samples. When looking at the beta diversity partitioning calculated from the OTUs/haplotype data between sites and seasons, turnover came over as the most important component (Fig. 5, Suppl. material 2: Fig. S3). On the contrary, the contribution of nestedness varied across seasons. In December, nestedness was a consistent component for St6 (Fig. 5c) and for St4 and St5 to a lesser extent while in July it was generally negligible (Fig. 5d, Suppl. material 2: Fig. S3). The correlation of nestedness at different taxonomic levels decreased with increasing the distance amongst them (Table 4). Regarding the

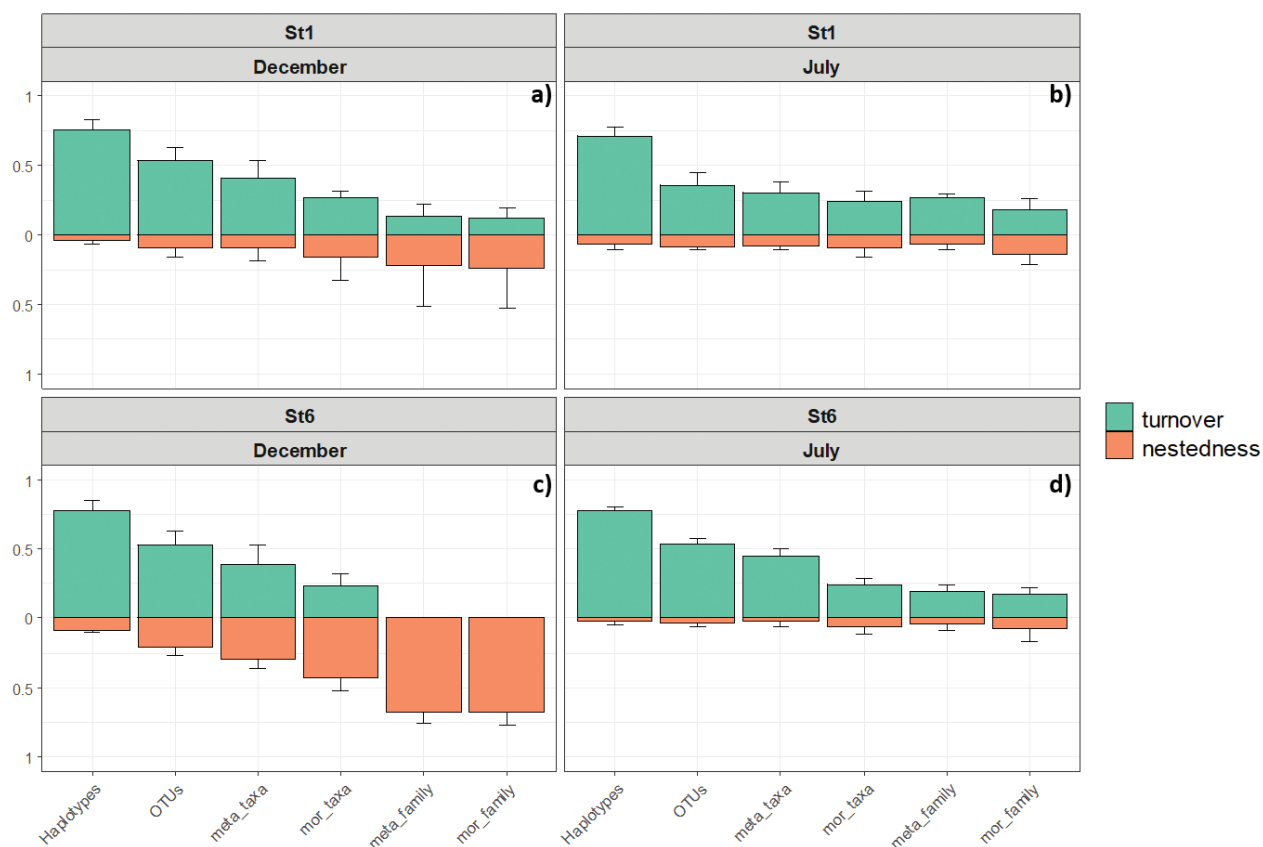


Figure 5. Values of the turnover and nestedness components for St1 and St6 in both December and July and for different identification levels (family, taxa = mixed taxonomic level) and methods (mor = morphological identification, meta = metabarcoding). The height of bars represents the average of the total dissimilarity between the target site and other sites, while the error bars the standard deviation.

Table 4. Spearman's rank correlation coefficients amongst turnover/nestedness fraction of beta diversity at different identification levels (fam = family, taxa = mixed taxonomic level) for both methods (mor = morphological identification, meta = metabarcoding).

fam mor	0.89	0.91	0.86	0.64	0.37
**	fam meta	0.87	0.76	0.48	0.29
**	**	taxa mor	0.86	0.61	0.35
**	**	**	taxa meta	0.84	0.53
**	*	**	**	OTU	0.59
*	*	*	**	**	haplotype

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, . $p < 0.1$

metabarcoding-generated data, good rank correlations were found between OTUs and the mixed taxonomic level (0.84, p -value < 0.01) and between the mixed taxonomic and family levels (0.76, p -value < 0.01), while the other correlations were always below 0.6 despite being significant. The inclusion of phylogenetic structure into the beta diversity partitioning analysis increased the amount of nestedness at the expenses of turnover (Suppl. material 2: Figs S4, S5).

Discussion

Comparison between morphology and metabarcoding

The agreement between morphological and metabarcoding identifications is an important step towards using metabarcoding in community ecology. In our work, metabarcoding detected most of the taxa identified with morphological identification. Similar results were found by Elbrecht et al. (2017), Emilson et al. (2017) and Kuntke et al. (2020) for freshwater macroinvertebrates. The mismatches recorded between both methods may be due to the lack of reference sequences and misclassifications in BOLD and the primer bias. The lack of reference sequences in BOLD probably had the strongest impact on the results of this study because 29 OTUs were assigned to higher taxonomic levels (e.g. order or class). Taxa with no reference database entry cannot be assigned to species with molecular methods. The incompleteness and lack of accuracy in the reference databases have been advocated as the most important problem leading to mismatches between morphological and molecular method results (Carew et al. 2017, Pawlowski et al.

2018, McGee et al. 2019). Similar problems were found by Andújar et al. (2018) and Elbrecht et al. (2017) for macroinvertebrates. Incomplete reference databases have also been reported by several authors for different biotic groups, such as littoral hard bottom communities (Wangenstein et al. 2018), macrophytes (Kuzmina et al. 2018) and ichthyoplankton (Hubert et al. 2015). In our work, we likely deflated the rate of families detected by metabarcoding because we did not trust hits below 90% similarity, a very conservative approach in terms of assignment. The problem of unassigned taxa is expected to improve in the future due to ongoing barcoding activities and because molecular methods are being increasingly used in ecological studies (Curry et al. 2018, Weigand et al. 2019).

Lack of reference sequences from BOLD is more likely for taxa with complex diagnostic characters (e.g. Hydracarina or Turbellaria). However, our results showed that some well-studied insect taxa were not assigned to their respective higher taxonomic level. In this context, wider divergence, i.e. larger than 10%, is reported within mayfly genera (Lucentini et al. 2011, Morinière et al. 2017). Therefore, the presence of morphologically cryptic taxa with great phylogenetic divergence and the use of fixed thresholds to define taxonomic levels (Brown et al. 2015), could explain our findings. In our work, some morphologically identified taxa included potential cryptic species (Jackson et al. 2014), as seen for instance for the mayfly *O. rhenana*. This species, the only Oligoneuriidae species reported for Italy, was amongst the most abundant ephemeropteran in our samples in July. However, this was detected only by metabarcoding at the family (OTU_112) and order (OTU_32) levels. The barcoding of *O. rhenana* individuals confirmed the metabarcoding results, indicating the presence of a potential cryptic species. Furthermore, the employed fixed threshold could have impeded the assignment of *Oligoneuriella* OTUs at the species or genus level as the mean intra- and interspecies distances of the *O. rhenana* sequences retrieved from BOLD were 12% and 15%, respectively. The herein applied use of a 95% threshold to assign genera and of 90% to assign families might not suffice, at least for *Oligoneuriella/Oligoneuriidae*. Despite more research being necessary to define the taxonomic status of this taxon (e.g. by analysing nuclear markers, morphological identification based on adult specimens, comparisons to morphology and genetic markers of type material), the implications of these findings are worth discussing. Metabarcoding can in fact help to identify inconsistencies in the current morphological taxonomy and to improve our freshwater fauna knowledge (Cardoni et al. 2015). Detecting cryptic species is also important, because they may have different environmental requirements (Eisenring et al. 2016, Leys et al. 2016), contrasting responses to multiple stressors (Sturmbauer et al. 1999, Macher et al. 2016, Beermann et al. 2018) and might provide valuable information on biogeographical processes.

Community structure detection with metabarcoding

We demonstrated that the increased taxonomic resolution provided by metabarcoding, even down to intraspecific diversity (haplotypes), is useful for interpreting macroinvertebrate community patterns as highlighted by alpha diversity, nMDS and, in particular, turnover/nestedness results. We found good concordance amongst the alpha diversity estimates for both the morphology and metabarcoding identification methods at the family and mixed levels (down to species, whenever possible). Concordance declined when comparing OTUs, or even haplotypes, with higher taxonomic levels. In this context, we can expect good concordance between different identification levels when the lower to higher taxonomic level ratio is low (Gaston 2000a). This happens, for example, when the number of species within each family is small and thus the ratio approaches 1. The use of molecular data is, in turn, expected to exacerbate this situation given the increased ability to assess biodiversity as highlighted in our work, where the number of entities to be studied (richness) increased by a factor of 10 when moving from the family level to OTUs/haplotypes.

Our results showed that the correlation amongst the ordinations performed with the data at different Linnean taxonomic levels for both morphology and metabarcoding was high (> 0.88). Good concordance between community patterns at different taxonomic levels has been found for arthropods (Timms et al. 2013), phytoplankton (Carneiro et al. 2010) and spiders (Cardoso et al. 2004). For macroinvertebrates, Slimani et al. (2019) reported a marked concordance of the community structure at the family and genus levels with those obtained at species level for different insect orders. Ordinations were also consistent when comparing abundance and presence-absence data, despite minor differences arising in ecological interpretations. In our work, the good correlation between presence-absence and abundance ordination was likely due to the transformation of the abundances made prior to running the nMDS. Transformation of abundances is the default option in the vegan package when wide abundance variability is present and is generally performed to reduce the effects of abundant taxa on the ordination result. In line with this, Heino (2008) found good concordance between abundance and presence-absence matrices only after data transformation. Despite the good concordance with the Linnaean taxonomic levels, the ordinations performed with OTUs/haplotypes were effective in tracking temporal and spatial trends in community structure and in detecting the intermittency effect in December. The clearer clustering of temporal trends with metabarcoding was probably due to the increased taxonomic resolution of OTUs, which led to more taxa being exclusive to only the first or second season.

Morphological identification and metabarcoding gave contrasting responses when looking at beta diver-

sity partitioning. This was especially true in December, when turnover turned out to be the most important fraction for OTU/haplotype level, while higher taxonomic levels highlighted the opposite pattern. In our work, not all organisms were identified at species level by either morphological identification or metabarcoding. This is a potential limitation of our study, as beta diversity dynamics at small spatial scale should be better highlighted by using finer taxonomic levels. The use of OTUs or inferred haplotypes allowed us to circumvent this problem by disengaging ecological patterns from the Linnean taxonomy. In this context, recent findings have shown that inferences in beta diversity dynamics based on genetic methods, performed better than those based on morphological identifications (Gill et al. 2012, Vodá et al. 2015, Bringlee et al. 2016). Including the phylogenetic information into the beta diversity partitioning analysis led to a decrease of the turnover percentage, meaning that OTUs and haplotypes found at intermittent sites were phylogenetically related to those of perennial sites. However, the inclusion of phylogenetic information lacks a direct link with desiccation, because strictly related species or genera can have different responses to drying (Suemoto et al. 2004, Gusev et al. 2014).

The increasing availability of data at higher taxonomic resolutions poses questions about the taxonomic level needed to study key ecological processes. At first, our results showed how an increased resolution was beneficial for the most abundant orders found with morphological identification. This was particularly evident for haplotypes, to which Diptera and Ephemeroptera contributed 82%. Haplotype richness could thus reflect both the abundance of organisms of a target group, the presence of bottleneck or founder effects, as well as the species-specific variability of the COI fragment tested in this study. Moreover, the use of OTUs or haplotypes instead of the family level data resulted in different beta diversity partitioning results. However, a finer taxonomic resolution might not always be beneficial for inferring ecological patterns from monitoring data. For example, Martin et al. (2016) demonstrated that the explanatory power of metacommunity models increased with coarse taxonomic resolution because of the assumed ecological interchangeability of species within genera or families. Moreover, the increase in turnover at the expense of nestedness when finer taxonomic levels were used could be related to the better chance of finding rarer taxa when increasing alpha diversity, rather than to structural patterns governed by flow intermittency. Additionally, the effect of a low sampling effort increases with increasing taxonomic resolution, which advises caution when interpreting nestedness/turnover data based on haplotypes inferred from very few individuals. Thus, the ideal taxonomic level for studying macroinvertebrate dynamics in river networks obviously depends on the research question. However, we argue that emerging genetic techniques have the potential to allow a more complete description of the

biological connectivity amongst sites, which will refine our knowledge of species-specific autecological preferences. The advantage of genetic data is that a lower resolution can be easily obtained by lumping taxa into higher ranks (family, order etc.).

Ecological patterns inferred from metabarcoding data

Recolonisation patterns of macroinvertebrates in IRES are driven by different factors, amongst which distance of a given intermittent site from the nearest perennial site and length and predictability of the dry period are the most important (Datry et al. 2017). After a dry phase, reduced alpha diversity at different taxonomic levels was evidenced for intermittent sites (Fig. 2). This trend was not observed in July, which denotes the resilience of the macroinvertebrate community to lack of water. Similarly to our findings, the recovery of the macroinvertebrate community to pre-drought levels was found by Skoulikidis et al. (2011) in an artificial intermittent stream in Greece. Datry (2012) reported similar results in the Albarine River, although the degree of recovery depended on the distance from the nearest perennial source.

nMDS results supported conclusions drawn from alpha diversity patterns, which revealed greater dispersion for December (post-drying) samples than for July (pre-drying) ones. The greater dispersion of December samples was driven by intermittent sites, which could host a nested subset of communities at perennial sites or a completely different pool of taxa. Community nestedness is expected to be dominant along flow intermittency gradients in IRES (Datry et al. 2014). In this context, our results revealed the opposite pattern for both sampling seasons. Given the distance from the nearest perennial source and the absence of supra-seasonal dry periods, the observed pattern was expected in July as 9 months had elapsed since flow resumption, which thus allowed complete recolonisation or emergence from riverbed by macroinvertebrates at intermittent sites. Recovery of the macroinvertebrate community to the pre-drying conditions is relatively fast in IRES (Acuña et al. 2005, Datry 2012). On the contrary, the predominance of turnover over nestedness upon rewetting was not expected and was clearly detectable only with the metabarcoding data. Recolonisation after rewetting was herein driven by a mixture of physiological adaptations to drying, dispersal abilities and behavioural adaptations. Organisms with desiccation resistance stages or with excellent dispersal abilities are favoured when the distance between perennial and intermittent sites is long, whereas dispersal of organisms is not limited when they are in close proximity (Bogan et al. 2013). Dispersal limitation was likely a minor factor in our study because of the short intermittent stretch and the maximum distance of 2.5 km to the nearest perennial source. After rewetting, however, 26% of OTUs were exclusive to intermittent sites, which suggest recolonisation from the Po River or from organisms with physiological or behavioural adaptations to intermittency. Intermittent sites hosted 16 exclusive OTUs,

with eight OTUs of Chironomidae, three OTUs of Simuliidae, one OTU of Trombidiformes, one OTU of *Calopteryx splendens* and one OTU of *Cloeon dipterum*. Chironomidae can survive dry periods in desiccation-resistant stages (Frouz et al. 2003) or by using the hyporheic zone as a refugium (Stubington and Datry 2013). Amongst the Chironomidae species exclusive for intermittent sites after re-wetting, *Cricotopus bicinctus* has often been found in both lentic and lotic temporary water bodies (Boix et al. 2001). Simuliidae larvae and their eggs are generally sensitive to desiccation (Adler and McCreddie 2002) and the presence of OTUs exclusive of intermittent sites is probably due to either the dispersal of adults and larvae or their resistance in the hyporheic zone (Lencioni and Spitale 2015). However, see Bogan et al. (2013) for exceptions. Trombidiformes present resistant stages, as rehydration experiments have clearly shown (Stubington and Datry 2013), but also a high drift propensity (Imbert and Perry 2000, Fenoglio et al. 2004). As *C. splendens* is not resistant to desiccation, its presence at intermittent sites was probably due to drift from upstream sites or from eggs hatched from late oviposition (Gallesi and Sacchi 2019). *C. dipterum* was present only at site 6 in December. Thus, its presence could be attributed to the spatial proximity of this site to the Po River confluence or to other perennial water sources in the nearby surroundings (e.g. drainage channels). *C. dipterum* is ovoviviparous and lays eggs containing mature embryos that hatch immediately after oviposition, which thus allows impacted sites to be rapidly recolonised (Degrange 1959).

Concluding remarks

Despite studying only few sites and limited repeated sampling over time, our work provided interesting insights into the use of metabarcoding to study ecological processes in IRES. At first, community metabarcoding yielded reliable information about taxonomic composition at all sites. In the Trebbia River, community dynamics inferred from metabarcoding data can be used alternatively to those derived from morphological identification when family/mixed identification levels and presence–absence information are needed. Second, we demonstrated that an increase in the identification level only moderately affects alpha diversity and community structure patterns, while it strongly affects beta diversity partitioning. The latter showed a switch from the nestedness to the turnover component moving from family to OTUs/haplotypes. Finally, an increased taxonomic resolution to species level using metabarcoding data was beneficial to interpret the observed patterns. Our work highlights how inferences drawn from highly resolved identification levels, even down to haplotypes, can provide a more comprehensive picture about the responses of biological communities to flow intermittency. This is a key point for IRES, where detailed information about recolonisation processes are needed to detect ongoing dynamics and for planning optimal conservation strategies.

Acknowledgements

A.L. was supported by a grant for a short-term scientific mission to the University of Duisburg-Essen, Germany, as part of COST Action CA15113 (SMIRES, Science and Management of Intermittent Rivers and Ephemeral Streams, www.smires.eu). This work was partially funded by Project PRIN-NOACQUA: responses of communities and ecosystem processes in intermittent rivers (Prot. 201572HW8F) and by the Consorzio di Bonifica di Piacenza. F.L., A.J.B., V.E. and V.M.A.Z. are members of and supported by COST Action DNAqua-Net (CA15219). The authors would like to thank Daniele Nizzoli, Debora Pucci, Beatrice Palmia and Edoardo Severini for their help during the fieldwork and Marie-Thérèse Werner and the Leeselab group for their help with the laboratory analysis. The authors also thank Simon Vitecek for his help in confirming the misclassification of *Rhyacophila pusilla* in BOLD.

References

- Acuña V, Muñoz I, Giorgi A, Omella M, Sabater F, Sabater S (2005) Drought and postdrought recovery cycles in an intermittent Mediterranean stream: structural and functional aspects. *Journal of the North American Benthological Society* 24: 919–933. <https://doi.org/10.1899/04-078.1>
- Adler PH, McCreddie JW (2002) Black Flies (Simuliidae). *Medical and Veterinary Entomology*. Academic Press. <https://doi.org/10.1016/B978-012510451-7/50013-X>
- 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>
- Baselga A (2010) Partitioning the turnover and nestedness components of beta diversity. *Global Ecology and Biogeography* 19: 134–143. <https://doi.org/10.1111/j.1466-8238.2009.00490.x>
- Baselga A, Orme D, Villeger S, Bortoli JD, Leprieur F, Logez M, Henriques-Silva R (2018) betapart: Partitioning beta diversity into turnover and nestedness components. <https://CRAN.R-project.org/package=betapart> [September 2, 2019]
- Beermann AJ, Zizka VMA, Elbrecht V, Baranov V, Leese F (2018) DNA metabarcoding reveals the complex and hidden responses of chironomids to multiple stressors. *Environmental Sciences Europe* 30: 26. <https://doi.org/10.1186/s12302-018-0157-x>
- Bogan MT, Boersma KS, Lytle DA (2013) Flow intermittency alters longitudinal patterns of invertebrate diversity and assemblage composition in an arid-land stream network. *Freshwater Biology* 58: 1016–1028. <https://doi.org/10.1111/fwb.12105>
- Boix D, Sala J, Moreno-Amich R (2001) The faunal composition of espolla pond (NE Iberian Peninsula): the neglected biodiversity of temporary waters. *Wetlands* 21: 577–592. [https://doi.org/10.1672/0277-5212\(2001\)021\[0577:TFCOEP\]2.0.CO;2](https://doi.org/10.1672/0277-5212(2001)021[0577:TFCOEP]2.0.CO;2)
- Bringloe TT, Cottenie K, Martin GK, Adamowicz SJ (2016) The importance of taxonomic resolution for additive beta diversity as revealed through DNA barcoding. *Genome* 59: 1130–1140. <https://doi.org/10.1139/gen-2016-0080>

- Brown EA, Chain FJJ, Crease TJ, MacIsaac HJ, Cristescu ME (2015) Divergence thresholds and divergent biodiversity estimates: can metabarcoding reliably describe zooplankton communities? *Ecology and Evolution* 5: 2234–2251. <https://doi.org/10.1002/ece3.1485>
- Bryant JA, Lamanna C, Morlon H, Kerkhoff AJ, Enquist BJ, Green JL (2008) Microbes on mountainsides: Contrasting elevational patterns of bacterial and plant diversity. *Proceedings of the National Academy of Sciences* 105: 11505–11511. <https://doi.org/10.1073/pnas.0801920105>
- Bush A, Compson ZG, Monk WA, Porter TM, Steeves R, Emilson E, Gagne N, Hajibabaei M, Roy M, Baird DJ (2019) Studying ecosystems with DNA metabarcoding: lessons from biomonitoring of aquatic macroinvertebrates. *Frontiers in Ecology and Evolution* 7. <https://doi.org/10.3389/fevo.2019.00434>
- Campaioli S, Ghetti PF, Minelli A, Ruffo S (1994) Manuale per il riconoscimento dei macroinvertebrati delle acque dolci italiane, Volume 1. Provincia Autonoma di Trento, 357 pp.
- Campaioli S, Ghetti PF, Minelli A, Ruffo S (1999) Manuale per il riconoscimento dei macroinvertebrati delle acque dolci italiane, Volume 2. Provincia Autonoma di Trento, 126 pp.
- Cardoni S, Tenchini R, Ficulle I, Piredda R, Simeone MC, Belfiore C (2015) DNA barcode assessment of Mediterranean mayflies (Ephemeroptera), benchmark data for a regional reference library for rapid biomonitoring of freshwaters. *Biochemical Systematics and Ecology* 62: 36–50. <https://doi.org/10.1016/j.bse.2015.07.035>
- Cardoso P, Silva I, de Oliveira NG, Serrano ARM (2004) Higher taxa surrogates of spider (Araneae) diversity and their efficiency in conservation. *Biological Conservation* 117: 453–459. <https://doi.org/10.1016/j.biocon.2003.08.013>
- Carew ME, Nichols SJ, Batovska J, Clair RS, Murphy NP, Blacket MJ, Shackleton ME (2017) A DNA barcode database of Australia's freshwater macroinvertebrate fauna. *Marine and Freshwater Research* 68: 1788–1802. <https://doi.org/10.1071/MF16304>
- Carneiro FM, Bini LM, Rodrigues LC (2010) Influence of taxonomic and numerical resolution on the analysis of temporal changes in phytoplankton communities. *Ecological Indicators* 10: 249–255. <https://doi.org/10.1016/j.ecolind.2009.05.004>
- Cuffney TF, Bilger MD, Haigler AM (2007) Ambiguous taxa: effects on the characterization and interpretation of invertebrate assemblages. *Journal of the North American Benthological Society* 26: 286–307. [https://doi.org/10.1899/0887-3593\(2007\)26\[286:A-TEOTC\]2.0.CO;2](https://doi.org/10.1899/0887-3593(2007)26[286:A-TEOTC]2.0.CO;2)
- Curry CJ, Gibson JF, Shokralla S, Hajibabaei M, Baird DJ (2018) Identifying North American freshwater invertebrates using DNA barcodes: are existing COI sequence libraries fit for purpose? *Freshwater Science* 37: 178–189. <https://doi.org/10.1086/696613>
- Datry T (2012) Benthic and hyporheic invertebrate assemblages along a flow intermittence gradient: effects of duration of dry events. *Freshwater Biology* 57: 563–574. <https://doi.org/10.1111/j.1365-2427.2011.02725.x>
- Datry T, Bonada N, Heino J (2016) Towards understanding the organisation of metacommunities in highly dynamic ecological systems. *Oikos* 125: 149–159. <https://doi.org/10.1111/oik.02922>
- Datry T, Bonada N, Boulton AJ (2017) Intermittent rivers and ephemeral streams: ecology and management. Academic Press, 624 pp.
- Datry T, Larned ST, Fritz KM, Bogan MT, Wood PJ, Meyer EI, Santos AN (2014) Broad-scale patterns of invertebrate richness and community composition in temporary rivers: effects of flow intermittence. *Ecography* 37: 94–104. <https://doi.org/10.1111/j.1600-0587.2013.00287.x>
- Degrange C (1959) L'ovolarviparité de *Cloëon dipterum* (L.) [Ephem. Baetidae]. *Bulletin de la Société entomologique de France* 64: 94–100.
- Döll P, Schmied HM (2012) How is the impact of climate change on river flow regimes related to the impact on mean annual runoff? A global-scale analysis. *Environmental Research Letters* 7: 014037. <https://doi.org/10.1088/1748-9326/7/1/014037>
- Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* 32: 1792–1797. <https://doi.org/10.1093/nar/gkh340>
- Edgar RC (2010) Search and clustering orders of magnitude faster than BLAST. *Bioinformatics (Oxford, England)* 26: 2460–2461. <https://doi.org/10.1093/bioinformatics/btq461>
- Edgar RC, Flyvbjerg H (2015) Error filtering, pair assembly and error correction for next-generation sequencing reads. *Bioinformatics* 31: 3476–3482. <https://doi.org/10.1093/bioinformatics/btv401>
- Eisenring M, Altermatt F, Westram AM, Jokela J (2016) Habitat requirements and ecological niche of two cryptic amphipod species at landscape and local scales. *Ecosphere* 7: e01319. <https://doi.org/10.1002/ecs2.1319>
- Eibrecht 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>
- Eibrecht V, Leese F (2017) Validation and development of COI metabarcoding primers for freshwater macroinvertebrate bioassessment. *Frontiers in Environmental Science* 5. <https://doi.org/10.3389/fevs.2017.00038>
- Eibrecht V, Steinke D (2019) Scaling up DNA metabarcoding for freshwater macrozoobenthos monitoring. *Freshwater Biology* 64: 380–387. <https://doi.org/10.7287/peerj.preprints.3456>
- Eibrecht V, Vamos EE, Steinke D, Leese F (2018) Estimating intraspecific genetic diversity from community DNA metabarcoding data. *PeerJ* 6: e4644. <https://doi.org/10.7717/peerj.4644>
- Eibrecht V, Vamos EE, 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>
- Emilson CE, Thompson DG, Venier LA, Porter TM, Swystun T, Chartrand D, Capell S, Hajibabaei M (2017) DNA metabarcoding and morphological macroinvertebrate metrics reveal the same changes in boreal watersheds across an environmental gradient. *Scientific Reports* 7: 12777. <https://doi.org/10.1038/s41598-017-13157-x>
- Fenoglio S, Bo T, Gallina G, Cucco M (2004) Vertical distribution in the water column of drifting stream macroinvertebrates. *Journal of Freshwater Ecology* 19: 485–492. <https://doi.org/10.1080/0270506.0.2004.9664923>
- Frouz J, Matena J, Ali A (2003) Survival strategies of chironomids (Diptera: Chironomidae) living in temporary habitats: a review. *European Journal of Entomology* 100: 459–465. <https://doi.org/10.14411/eje.2003.069>
- Galesi MM, Sacchi R (2019) Voltinism and larval population structure of *Calopteryx splendens* (Odonata: Calopterygidae) in the Po Valley. *International Journal of Odonatology* 22: 21–30. <https://doi.org/10.1080/13887890.2018.1563916>

- Gaston KJ (2000a) Biodiversity: higher taxon richness. *Progress in Physical Geography* 24: 11. <https://doi.org/10.1177/030913330002400108>
- Gaston KJ (2000b) Global patterns in biodiversity. *Nature* 405: 220–227. <https://doi.org/10.1038/35012228>
- 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>
- Gill BA, Harrington RA, Kondratieff BC, Zamudio KR, Poff NL, Funk WC (2012) Morphological taxonomy, DNA barcoding, and species diversity in southern Rocky Mountain headwater streams. *Freshwater Science* 33: 288–301. <https://doi.org/10.1086/674526>
- Gusev O, Suetsugu Y, Cornette R, Kawashima T, Logacheva MD, Kondrashov AS, Penin AA, Hatanaka R, Kikuta S, Shimura S, Kanamori H, Katayose Y, Matsumoto T, Shagimardanova E, Alexeev D, Govorun V, Wisecaver J, Mikheyev A, Koyanagi R, Fujie M, Nishiyama T, Shigenobu S, Shibata TF, Golygina V, Hasebe M, Okuda T, Satoh N, Kikawada T (2014) Comparative genome sequencing reveals genomic signature of extreme desiccation tolerance in the anhydrobiotic midge. *Nature Communications* 5: 4784. <https://doi.org/10.1038/ncomms5784>
- Hajibabaei M, Shokralla S, Zhou X, Singer GAC, Baird DJ (2011) Environmental barcoding: a next-generation sequencing approach for biomonitoring applications using river benthos. *PLOS ONE* 6: e17497. <https://doi.org/10.1371/journal.pone.0017497>
- Heibl C (2008) PHYLOCH: R language tree plotting tools and interfaces to diverse phylogenetic software packages.
- Heino J (2008) Influence of taxonomic resolution and data transformation on biotic matrix concordance and assemblage–environment relationships in stream macroinvertebrates. *Boreal Environmental Research* 13: 359–369.
- Hubert N, Espiau B, Meyer C, Planes S (2015) Identifying the ichthyoplankton of a coral reef using DNA barcodes. *Molecular Ecology Resources* 15: 57–67. <https://doi.org/10.1111/1755-0998.12293>
- Imbert JB, Perry JA (2000) Drift and benthic invertebrate responses to stepwise and abrupt increases in non-scouring flow. *Hydrobiologia* 436: 191–208. <https://doi.org/10.1023/A:1026582218786>
- Jackson JK, Battle JM, White BP, Pilgrim EM, Stein ED, Miller PE, Sweeney BW (2014) Cryptic biodiversity in streams: a comparison of macroinvertebrate communities based on morphological and DNA barcode identifications. *Freshwater Science* 33: 312–324. <https://doi.org/10.1086/675225>
- Jones FC (2008) Taxonomic sufficiency: The influence of taxonomic resolution on freshwater bioassessments using benthic macroinvertebrates. *Environmental Reviews* 16: 45–69. <https://doi.org/10.1139/A07-010>
- Knouft JH, Losos JB, Gior RE, Kolbe JJ (2006) Phylogenetic analysis of the evolution of the niche in lizards of the Anolis Sagrei Group. *Ecology* 87: S29–S38. [https://doi.org/10.1890/0012-9658\(2006\)87\[29:PAOTEO\]2.0.CO;2](https://doi.org/10.1890/0012-9658(2006)87[29:PAOTEO]2.0.CO;2)
- Kuntke F, de Jonge N, Hesselsoe M, Lund Nielsen J (2020) Stream water quality assessment by metabarcoding of invertebrates. *Ecological Indicators* 111: 105982. <https://doi.org/10.1016/j.ecolind.2019.105982>
- Kuzmina ML, Braukmann TWA, Zakharov EV (2018) Finding the pond through the weeds: eDNA reveals underestimated diversity of pondweeds. *Applications in Plant Sciences* 6. <https://doi.org/10.1002/aps3.1155>
- Lake PS (2003) Ecological effects of perturbation by drought in flowing waters. *Freshwater Biology* 48: 1161–1172. <https://doi.org/10.1046/j.1365-2427.2003.01086.x>
- Leese F, Altermatt F, Bouchez A, Ekrem T, Hering D, Meissner K, Mergen P, Pawlowski J, Piggott J, Rimet F, Steinke D, Taberlet P, Weigand A, Abarenkov K, Beja P, Bervoets L, Björnsdóttir S, Boets P, Boggero A, Bones A, Borja Á, Bruce K, Bursić V, Carlsson J, Čiampor F, Čiamporová-Zatovičová Z, Coissac E, Costa F, Costache M, Creer S, Csabai Z, Deiner K, DelValls Á, Drakare S, Duarte S, El-eršek T, Fazi S, Fišer C, Flot J-F, Fonseca V, Fontaneto D, Grabowski M, Graf W, Guðbrandsson J, Hellström M, Hershkovitz Y, Hollingsworth P, Japoshvili B, Jones J, Kahlert M, Stroil BK, Kasapidis P, Kelly M, Kelly-Quinn M, Keskin E, Kõljalg U, Ljubešić Z, Maček I, Mächler E, Mahon A, Marečková M, Mejdzancic M, Mircheva G, Montagna M, Moritz C, Mulk V, Naumoski A, Navodaru I, Padisák J, Pálsson S, Panksep K, Penev L, Petrussek A, Pfannkuchen M, Primmer C, Rinkevich B, Rotter A, Schmidt-Kloiber A, Segurado P, Speksnijder A, Stoev P, Strand M, Šulčius S, Sundberg P, Traugott M, Tsigenopoulos C, Turon X, Valentini A, Hoom B van der, Várbiro G, Hadjilyra MV, Viguri J, Vitonyté I, Vogler A, Vrålstad T, Wägele W, Wenne R, Winding A, Woodward G, Zegura B, 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>
- Lencioni V, Spitale D (2015) Diversity and distribution of benthic and hyporheic fauna in different stream types on an alpine glacial floodplain. *Hydrobiologia* 751: 73–87. <https://doi.org/10.1007/s10750-014-2172-2>
- Lencioni V, Marziali L, Rossaro B (2012) Chironomids as bioindicators of environmental quality in mountain springs. *Freshwater Science* 31: 525–541. <https://doi.org/10.1899/11-038.1>
- Leprieur F, Albouy C, Bortoli JD, Cowman PF, Bellwood DR, Mouillot D (2012) Quantifying phylogenetic beta diversity: distinguishing between “true” turnover of lineages and phylogenetic diversity gradients. *PLOS ONE* 7: e42760. <https://doi.org/10.1371/journal.pone.0042760>
- Leray M, Knowlton N (2017) Random sampling causes the low reproducibility of rare eukaryotic OTUs in Illumina COI metabarcoding. *Peer J* 5: e3006. <https://doi.org/10.7717/peerj.3006>
- Leys M, Keller I, Räsänen K, Gattolliat J-L, Robinson CT (2016) Distribution and population genetic variation of cryptic species of the Alpine mayfly *Baetis alpinus* (Ephemeroptera: Baetidae) in the Central Alps. *BMC Evolutionary Biology* 16: 77. <https://doi.org/10.1186/s12862-016-0643-y>
- Losos JB (2008) Phylogenetic niche conservatism, phylogenetic signal and the relationship between phylogenetic relatedness and ecological similarity among species. *Ecology Letters* 11: 995–1003. <https://doi.org/10.1111/j.1461-0248.2008.01229.x>
- Lucentini L, Rebora M, Puletti ME, Gliarelli L, Fontaneto D, Gaino E, Panara F (2011) Geographical and seasonal evidence of cryptic diversity in the *Baetis rhodani* complex (Ephemeroptera, Baetidae) revealed by means of DNA taxonomy. *Hydrobiologia* 673: 215–228. <https://doi.org/10.1007/s10750-011-0778-1>
- Macher JN, 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>

- Malicky H, Changhai S (2002) 25 new species of Rhyacophilidae (Trichoptera) from China. *Linzer biologische Beiträge* 34: 541–561.
- Martin GK, Adamowicz SJ, Cottenie K (2016) Taxonomic resolution based on DNA barcoding affects environmental signal in meta-community structure. *Freshwater Science* 35: 701–711. <https://doi.org/10.1086/686260>
- Martin M (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet journal* 17: 10–12. <https://doi.org/10.14806/ej.17.1.200>
- McGee KM, Robinson CV, Hajibabaei M (2019) Gaps in DNA-based biomonitoring across the globe. *Frontiers in Ecology and Evolution* 7. <https://doi.org/10.3389/fevo.2019.00337>
- Miller KB, Bergsten J (2016) *Diving Beetles of the World: Systematics and Biology of the Dytiscidae*. Johns Hopkins University Press, Baltimore, Maryland, 336 pp.
- Mittelbach GG, McGill BJ (2019) *Community Ecology*. 2^o edition. OUP Oxford, 432 pp. <https://doi.org/10.1093/oso/9780198835851.001.0001>
- Morinière J, Hendrich L, Balke M, Beermann AJ, König T, Hess M, Koch S, Müller R, Leese F, Hebert PDN, Hausmann A, Schubart CD, Haszprunar G (2017) A DNA barcode library for Germany's mayflies, stoneflies and caddisflies (Ephemeroptera, Plecoptera and Trichoptera). *Molecular Ecology Resources* 17: 1293–1307. <https://doi.org/10.1111/1755-0998.12683>
- Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGinn D, Minchin PR, O'Hara RB, Simpson GL, Solymos P, Stevens MHH, Szoecs E, Wagner H (2019) *vegan: Community Ecology Package*. <https://CRAN.R-project.org/package=vegan> [May 9, 2019]
- Pawlowski J, Kelly-Quinn M, Altermatt F, Apothéloz-Perret-Gentil L, Beja P, Boggero A, Borja A, Bouchez A, Cordier T, Domaizon I, Feio MJ, Filipe AF, Fornaroli R, Graf W, Herder J, van der Hoorn B, Iwan Jones J, Sagova-Mareckova M, Moritz C, Barquin J, Piggott JJ, Pinna M, Rimet F, Rinkevich B, Sousa-Santos C, Specchia V, Trobajo R, Vasselon V, Vitecek S, Zimmerman J, Weigand A, Leese F, Kahlert M (2018) The future of biotic indices in the ecogenomic era: Integrating (e)DNA metabarcoding in biological assessment of aquatic ecosystems. *Science of the Total Environment* 637–638: 1295–1310. <https://doi.org/10.1016/j.scitotenv.2018.05.002>
- Peres-Neto PR, Jackson DA (2001) How well do multivariate data sets match? The advantages of a Procrustean superimposition approach over the Mantel test. *Oecologia* 129: 169–178. <https://doi.org/10.1007/s004420100720>
- R Core Team (2017) *A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. <http://CRAN.R-project.org/>
- 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>
- Resh VH, Unzicker JD (1975) Water quality monitoring and aquatic organisms: the importance of species identification. *Journal – Water Pollution Control Federation* 47: 9–19.
- Rognes T, Flouri T, Nichols B, Quince C, Mahé F (2016) VSEARCH: a versatile open source tool for metagenomics. *Peer J* 4: e2584. <https://doi.org/10.7717/peerj.2584>
- Skoulikidis NT, Vardakas L, Karaouzas I, Economou AN, Dimitriou E, Zogaris S (2011) Assessing water stress in Mediterranean lotic systems: insights from an artificially intermittent river in Greece. *Aquatic Sciences* 73: 581. <https://doi.org/10.1007/s00027-011-0228-1>
- Slimani N, Sánchez-Fernández D, Guilbert E, Boumaïza M, Guareschi S, Thioulouse J (2019) Assessing potential surrogates of macroinvertebrate diversity in North-African Mediterranean aquatic ecosystems. *Ecological Indicators* 101: 324–329. <https://doi.org/10.1016/j.ecolind.2019.01.017>
- Stamatakis A (2014) RAXML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics (Oxford, England)* 30: 1312–1313. <https://doi.org/10.1093/bioinformatics/btu033>
- Stubbington R, Detry T (2013) The macroinvertebrate seedbank promotes community persistence in temporary rivers across climate zones. *Freshwater Biology* 58: 1202–1220. <https://doi.org/10.1111/fwb.12121>
- Stubbington R, Chadd R, Cid N, Csabai Z, Miliša M, Morais M, Munné A, Pařil P, Pešić V, Tziortzis I, Verdonschot RCM, Detry T (2018) Biomonitoring of intermittent rivers and ephemeral streams in Europe: Current practice and priorities to enhance ecological status assessments. *Science of the Total Environment* 618: 1096–1113. <https://doi.org/10.1016/j.scitotenv.2017.09.137>
- Sturmbauer C, Opadiya GB, Niederstätter H, Riedmann A, Dallinger R (1999) Mitochondrial DNA reveals cryptic oligochaete species differing in cadmium resistance. *Molecular Biology and Evolution* 16: 967–974. <https://doi.org/10.1093/oxfordjournals.molbev.a026186>
- Suemoto T, Kawai K, Imabayashi H (2004) A comparison of desiccation tolerance among 12 species of chironomid larvae. *Hydrobiologia* 515: 107–114. <https://doi.org/10.1023/B:HYDR.0000027322.11005.20>
- 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>
- Tachet H (2010) *Invertébrés d'eau douce: Systématique, biologie, écologie*. édition revue et augmentée édition. CNRS EDITIONS, Paris, 608 pp.
- Timms LL, Bowden JJ, Summerville KS, Buddle CM (2013) Does species-level resolution matter? Taxonomic sufficiency in terrestrial arthropod biodiversity studies. *Insect Conservation and Diversity* 6: 453–462. <https://doi.org/10.1111/icad.12004>
- Vilmi A, Karjalainen SM, Nokela T, Tolonen K, Heino J (2016) Unravelling the drivers of aquatic communities using disparate organismal groups and different taxonomic levels. *Ecological Indicators* 60: 108–118. <https://doi.org/10.1016/j.ecolind.2015.06.023>
- Vodá R, Dapporto L, Dincă V, Vila R (2015) Cryptic matters: overlooked species generate most butterfly beta-diversity. *Ecography* 38: 405–409. <https://doi.org/10.1111/ecog.00762>
- Wallington TJ, Hobbs RJ, Moore SA (2005) Implications of current ecological thinking for biodiversity conservation: a review of the salient issues. *Ecology and Society* 10. <https://doi.org/10.5751/ES-01256-100115>
- 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. *Peer J* 6: e4705. <https://doi.org/10.7717/peerj.4705>
- Weigand H, Beermann AJ, Čiampor F, Costa FO, Csabai Z, Duarte S, Geiger MF, Grabowski M, Rimet F, Rulik B, Strand M, Szucsich N, Weigand AM, Willassen E, Wyler SA, Bouchez A, Borja A, Čiamporová-Zaťovičová Z, Ferreira S, Dijkstra K-DB, Eisendle U, Freyhof J, Gadawski P, Graf W, Haegerbaeumer A, van der Hoorn BB, Japoshvili B, Keresztes L, Keskin E, Leese F, Macher JN, Ma-

- mos T, Paz G, Pešić V, Pfannkuchen DM, Pfannkuchen MA, Price BW, Rinkevich B, Teixeira MAL, Várбірó G, Ekrem T (2019) DNA barcode reference libraries for the monitoring of aquatic biota in Europe: Gap-analysis and recommendations for future work. *Science of the Total Environment* 678: 499–524. <https://doi.org/10.1016/j.scitotenv.2019.04.247>
- Weiss M, Leese F (2016) Widely distributed and regionally isolated! Drivers of genetic structure in *Gammarus fossarum* in a human-impacted landscape. *BMC Evolutionary Biology* 16: 153. <https://doi.org/10.1186/s12862-016-0723-z>
- Wickham H (2009) *ggplot2: Elegant graphics for data analysis*. Springer-Verlag, New York. <https://doi.org/10.1007/978-0-387-98141-3>
- Zickovich JM, Bohonak AJ (2007) Dispersal ability and genetic structure in aquatic invertebrates: a comparative study in southern California streams and reservoirs. *Freshwater Biology* 52: 1982–1996. <https://doi.org/10.1111/j.1365-2427.2007.01822.x>
- Zizka VMA, Weiss M, Leese F (2020) Can metabarcoding resolve intraspecific genetic diversity changes to environmental stressors? A test case using river macrozoobenthos. *Metabarcoding and Metagenomics* 4: e51925. <https://doi.org/10.1101/2020.03.08.982561>

Supplementary material 1

Raw data

Authors: Alex Laini

Data type: raw data

Copyright notice: This dataset is made available under the Open Database License (<http://opendatacommons.org/licenses/odbl/1.0/>). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited. Link: <https://doi.org/10.3897/mbmg.4.51433.suppl1>

Supplementary material 2

Table S1, Figures S1–S5

Authors: Alex Laini

Data type: additional tables and figures

Copyright notice: This dataset is made available under the Open Database License (<http://opendatacommons.org/licenses/odbl/1.0/>). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited. Link: <https://doi.org/10.3897/mbmg.4.51433.suppl2>