Received: 7 October 2020 | Revised: 25 January 2021 | Accepted: 26 January 2021

DOI: 10.1111/1755-0998.13341

MOLECULAR ECOLOGY RESOURCES WILEY

RESOURCE ARTICLE

Two different approaches of microbial community structure characterization in riverine epilithic biofilms under multiple stressors conditions: Developing molecular indicators

Lorenzo Pin[1,2] | Alexander Eiler[2,3] | Stefano Fazi[4] | Nikolai Friberg[1,5,6]

[1]Norsk Institutt for Vannforskning (NIVA), Oslo, Norway

[2]Section for Aquatic Biology and Toxicology, Department of Biosciences, Centre for Biogeochemistry in the Anthropocene, University of Oslo, Norway

[3]eDNA solutions AB, Mölndal, Sweden

[4]Water Research Institute, IRSA-CNR, Monterotondo, Roma, Italy

[5]Freshwater Biological Section, University of Copenhagen, Copenhagen, Denmark

[6]School of Geography, University of Leeds, Leeds, UK

Correspondence
Lorenzo Pin, Norsk Institutt for Vannforskning (NIVA), Oslo, Norway.
Email: Lorenzo.pin@niva.no

Funding information
Marie Skłodowska-Curie grant Agreement (MSCA), Grant/Award Number: 765553. SF participation was supported by the CNR Short Term Mobility Programme 2019.

Abstract

Microbial communities are major players in the biogeochemical processes and ecosystem functioning of river networks. Despite their importance in the ecosystem, biomonitoring tools relying on prokaryotes are still lacking. Only a few studies have employed both metabarcoding and quantitative techniques such as catalysed reported deposition fluorescence *in situ* hybridization (CARD-FISH) to analyse prokaryotic communities of epilithic biofilms in river ecosystems. We intended to investigate the efficacy of both techniques in detecting changes in microbial community structure associated with environmental drivers. We report a significant correlation between the prokaryotic community composition and pH in rivers from two different geographical areas in Norway. Both CARD-FISH and metabarcoding data were following the pattern of the environmental variables, but the main feature distinguishing the community composition was the regional difference itself. Beta-dispersion analyses on both CARD-FISH abundance and metabarcoding data revealed higher accuracy of metabarcoding to differentiate regions and river systems. The CARD-FISH results showed high variability, even for samples within the same river, probably due to some unmeasured microscale ecological variability which we could not resolve. We also present a statistical method, which uses variation coefficient and overall prevalence of taxonomic groups, to detect possible biological indicators among prokaryotes using metabarcoding data. The development of new prokaryotic bioindicators would benefit from both techniques used in this study, but metabarcoding seems to be faster and more reliable than CARD-FISH for large scale bio-assessment.

KEYWORDS
biofilm, bioindicator, CARD-FISH, prokaryotes, river, sequencing

1 | INTRODUCTION

River systems are extremely dynamic ecosystems, providing a variety of services to humans (Arthington et al., 2018). Since ancient times rivers have been exploited in several ways: drinking water sources, water for agriculture, hydropower, cooling systems for industries, recreation, etc. (Poff, 2018; Vörösmarty et al., 2005). These activities have impaired natural flow variability, causing changes to habitats and the biodiversity of adjacent areas (Davies et al., 2014; Poff et al., 1997).

Biodiversity loss is one of the biggest concerns for river ecosystems, as most rivers are being exploited and increasingly losing species (Bunn & Arthington, 2002; Das Gupta, 2008; Merritt et al., 2010). Biodiversity loss and reduced abundance are becoming evident for several biological groups, such as insects, for which recent studies have shown global changes in freshwater (Baranov et al., 2020; Hallmann et al., 2017, 2020; Sánchez-Bayo & Wyckhuys, 2019).

In contrast, very little is known about how microbial diversity is influenced by anthropogenic stress. Microbial communities play a fundamental part in driving ecosystem processes and play a vital role at the base of riverine food webs (Demars et al., 2020; Robbins et al., 2017). Despite their small physical dimensions, microbes are key drivers of organic matter decomposition in fluvial ecosystems and mineralization of nutrients, making them available for higher levels of the riverine food web (Demars et al., 2020). This functional role is particularly true for the microbial communities living in epilithic biofilms with a high sediment to water phase ratio, producing a large extent of reactive surfaces in streams (Battin et al., 2008).

The role of microbial communities is today often only inferred in studies on ecosystem functioning, for example by using fine mesh bags in leaf litter decomposition studies (Gessner & Chauvet, 2002; Tiegs et al., 2019; Woodward et al., 2012). Mapping microbial diversity, and the functional traits related to key ecosystem processes, has huge potential in increasing our understanding of drivers of ecosystem functioning in the years to come (Besemer, 2015). Moreover, due to their high sensitivity to pollution and fast response to environmental changes, bacterial assemblages could complement the information provided by benthic metazoan communities as indicators of human-induced impacts, but this biological component has not yet been well explored in this regard (Caruso et al., 2016; Szabo et al., 2007). Jackson et al. (2016) strongly argued for use of sequencing of microbial communities as part of next-generation biomonitoring tools. In Europe there is currently a significant lack of prokaryotic indicators in the EU Water Framework Directive (2000/60/EC) and national and international legislations (Heiskanen et al., 2016), although all other relevant biological groups, including microalgae, are included in the assessment of freshwater status (Birk et al., 2012). Obvious reasons for the absence of prokaryotic communities in bioassessment have been the cost of taxonomic analyses and a lack of knowledge of their indicator value in terms of natural variability and human impacts. However, the easy access to high-throughput sequencing (HTS) technologies, allowing the quick taxonomic identification of bacterial assemblages, has led to several attempts to find prokaryotic indicators. Some of the most relevant studies focused on sediments of coastal areas and estuaries (Aylagas et al., 2017; Borja, 2018).

In freshwaters, a first attempt to use prokaryotes as bioindicators focused on quantitative techniques such as real-time qPCR selecting a few prokaryotic strains associated with specific chemicals and water quality parameters (Nzewi et al., 2009). Others have tried to address the lack of prokaryotic bioindicators for freshwater ecosystems, mainly using qPCR to quantitatively analyse specific functional genes (Thompson et al., 2016) and 16S amplicon sequencing to point out changes in the community structure of freshwater prokaryotes impacted by human activities (Salis et al., 2017; Simonin et al., 2019). Martínez-Santos et al. (2018) used both qPCR and 16S amplicon sequencing to analyse the effects of wastewater effluents, on structure and function of the prokaryotic communities dwelling in Deba river sediments. In our study, we focused on epilithic biofilm communities which, in addition to being of key importance for ecosystem functioning in rivers, have been shown to be more sensitive to water quality features compared to those dwelling on high organic matter (OM)-loaded substrates (leaves, roots, wood) (Fazi et al., 2005). Hence, they could provide reliable information regarding human and natural pressures on the environment. We explored the potential of two very different techniques as biodiversity indicators for prokaryotes in epilithic biofilms: a quantitative method, catalysed reported deposition fluorescence *in situ* Hybridization (CARD-FISH; Pernthaler et al., 2002); and a more qualitative method, metabarcoding of 16S rRNA amplicons (Johnson et al., 2019). The two techniques were used to characterize the microbial community structure of rivers in Norway, ranging from river systems affected by natural disturbances, to rivers affected by various types of human impacts, such as dams, hydropower developments and wastewater outlets. Two different geographical regions were selected, characterized by distinct geological and chemical features (Steinnes et al., 1993), probably influencing the prokaryotic community structure of epilithic biofilms.

We hypothesized that: (i) CARD-FISH and metabarcoding will provide similar patterns regarding the overall microbial community structure, but they will give different insights at different spatial scales (regional vs. microscale); (ii) that the community structure of epilithic biofilms would be influenced by both human perturbation and natural conditions such as the geological setting.

The advantage of metabarcoding to CARD-FISH is lower operating costs, which might make it better suited for use in modern biomonitoring networks if the methods yield comparable results in terms of describing prokaryotic communities. There is already evidence for the effectiveness of metabarcoding as a biomonitoring tool, for both eukaryotes and prokaryotes, with results comparable to the traditional methods based on species morphology (Cordier et al., 2019). Metabarcoding of prokaryotic communities would be complementary to the traditionally used bioindicators such as benthic macroinvertebrates, diatoms and fish, which today are the most commonly used group in impact assessment (e.g Birk et al., 2012;

Friberg, 2014). It would provide valuable insights into the black box of biodiversity in riverine ecosystems, namely the microbial communities which may be pivotal as early warning indicators of human and natural pressures (Besemer, 2015; Widder et al., 2014).

2 | MATERIALS AND METHODS

2.1 | Study design and samplings

Two different geographical regions in Norway were selected for the present study, characterized by distinct geological and chemical features, in particular in terms of acid neutralizing capacity (Steinnes et al., 1993) (Figure 1). A total of 16 sites was sampled, embracing an array of environmental gradients (acidity of water, geology, human impacts). In one area (in the Oslo region in southeast Norway), bedrock of the rivers is dominated by lower Palaeozoic sedimentary rocks (limestone and shales) (Calner, 2013). In the other area (Vest and Aust-Agder, southwest Norway) the bedrock mainly comprises magmatic and metamorphic rocks (Slagstad et al., 2018). The 16 sampling sites were situated in four catchments, two in each region: Lysakerelva and Sandvikselva in the Oslo region; and Arendalvassdraget and Mandalselva in the Agder region. Within these four catchments we selected four rivers impacted by hydropower and dams (Lysaker, Lomma, Nidelva and Mandal) and four free-flowing rivers (Iselva, Finnsåna, Haugedøl and Stigvasselva) which were our control sites (Table S1). For each control site, we selected only one sampling site, while for the impacted rivers we selected three sampling sites, one upstream from the dam or hydropower plant, so that this site could be comparable to our control sites as being virtually unimpacted. The second site was always set immediately below the dam, where we expected impact to be highest. The third site was further downstream, where the water was mixed, and the effect of the dam was not evident.

When establishing a sampling site, it was georeferenced using a global positioning system application (GPS Status version 2.0.4 (36), Maplewave Studio).

The sampling campaign was carried out in September 2018. Water samples were collected at each site approximately at 10 cm below the surface. A 500-ml water sample was taken at each site for chemical analysis by using polyethylene bottles. Contextually another 60-ml water sample was taken for the analysis of metals by using polyethylene bottles pretreated with a 1% HNO_3 solution. The samples were immediately placed at 4°C in the field and brought back to the laboratory for analysis.

Biofilm samples were collected at each site after taking the water sample. Five rocks (average individual surface 118 ± 22.4 cm^2) were randomly taken within a 50-m^2 area and brushed in the field with a sterile toothbrush to collect the epilithic biofilm. The biofilm brushed from the five rocks was pooled together to give a single sample for each site (16 biofilm samples in total). The pooled biofilm samples were suspended in 65 ml of ultrapure MilliQ water. An aliquot of each sample (50 ml), to be used for hybridization *in situ* (CARD-FISH), was added to 50 ml of pure ethanol to prevent ice formation and consequent cell lysis; the remaining 15 ml, to be used for DNA extraction, was placed in a 15-ml Falcon tube. We had 16 subsamples for CARD-FISH in total and 16 subsamples for sequencing in total. Both biofilm samples were kept cool at 4°C until arrival ae the laboratory where they were stored at −20°C until further processing.

2.2 | Physicochemical water parameters

Water temperature, pH and electrical conductivity (EC) were measured *in situ* with a multiparameter portable meter (WTW ProfiLine Multi 3320).

Water samples were analysed by NIVA (Norwegian Institute for Water Research, Oslo) for the following parameters: ammonium (NH_4), calcium (Ca^{2+}), magnesium (Mg^{2+}), sulphate (SO_4^{2-}) (NS-EN ISO 10304–1 [anions], NS -EN ISO 14911 [cations] [C4-4]), phosphate (PO_4^{3-}) (Mod. NS 4724 [D1-3]), Fe-tot, Fe (II), Fe (III) (intern method [EKSTERN_ALS]), dissolved organic carbon (DOC), total organic carbon (TOC) (NS-EN 1484:1997 [G4-2]), NO_2^-+ NO_3^- (Mod. NS 4745:1991 [D3-3]), total nitrogen (TN) (NIVA intern method), total phosphorus (TP) (Mod. NS 4725 [D2-1]). The quantification limits were: (NH_4^+) 2 μg L^{-1}; (Ca^{2+}, Mg^{2+}) 0.002 mg L^{-1}; (SO_4^{2-}) 0.005 mg L^{-1}; (PO_4^{3-}) 1 μg L^{-1}; (Fe-tot) 0.0020 mg L^{-1}; (Fe (II), Fe (III)) 0.01 mg L^{-1}; (TOC, DOC) 0.10 mg L^{-1}; (NO_2^-+ NO_3) 1 μg L^{-1}; (TN) 50 μg L^{-1}; (TP) 1 μg L^{-1}.

2.3 | Biofilm biomass quantification

For biomass quantification, we used the ash free dry mass (AFDW) content of the biofilm samples. Two replicates subaliquots (~1 g wet weight) were taken from each sample and preserved for hybridization *in situ*. The subaliquots were dried at 60°C in a thermostatic oven for 72 hr to obtain the dry weights. Subsequently the dried aliquots were pooled together and burned in a muffle oven at 550°C for 3 hr to obtain the ash weights. Subtracting the ash weights from the pooled dry weights we were able to measure the biomass content of our samples.

2.4 | Total prokaryotic abundance and single cell hybridization (CARD-FISH)

The total prokaryotic cell abundance was assessed by 4′-6-diamidino-2-phenylindole (DAPI, Vector Laboratories) staining, following extraction and detection procedures described in Amalfitano and Fazi (2008). Briefly, 1 g of biofilm (collected from the pellet of the 50-ml samples preserved in ethanol after being centrifuged [2795 G force for 10 min]) was fixed in formaldehyde solution (final concentration 2.0%), and amended with Tween 20 (final concentration 0.5%) and sodium pyrophosphate (1 g L^{-1} final concentration), resulting in 10 ml of solution containing biofilm. The

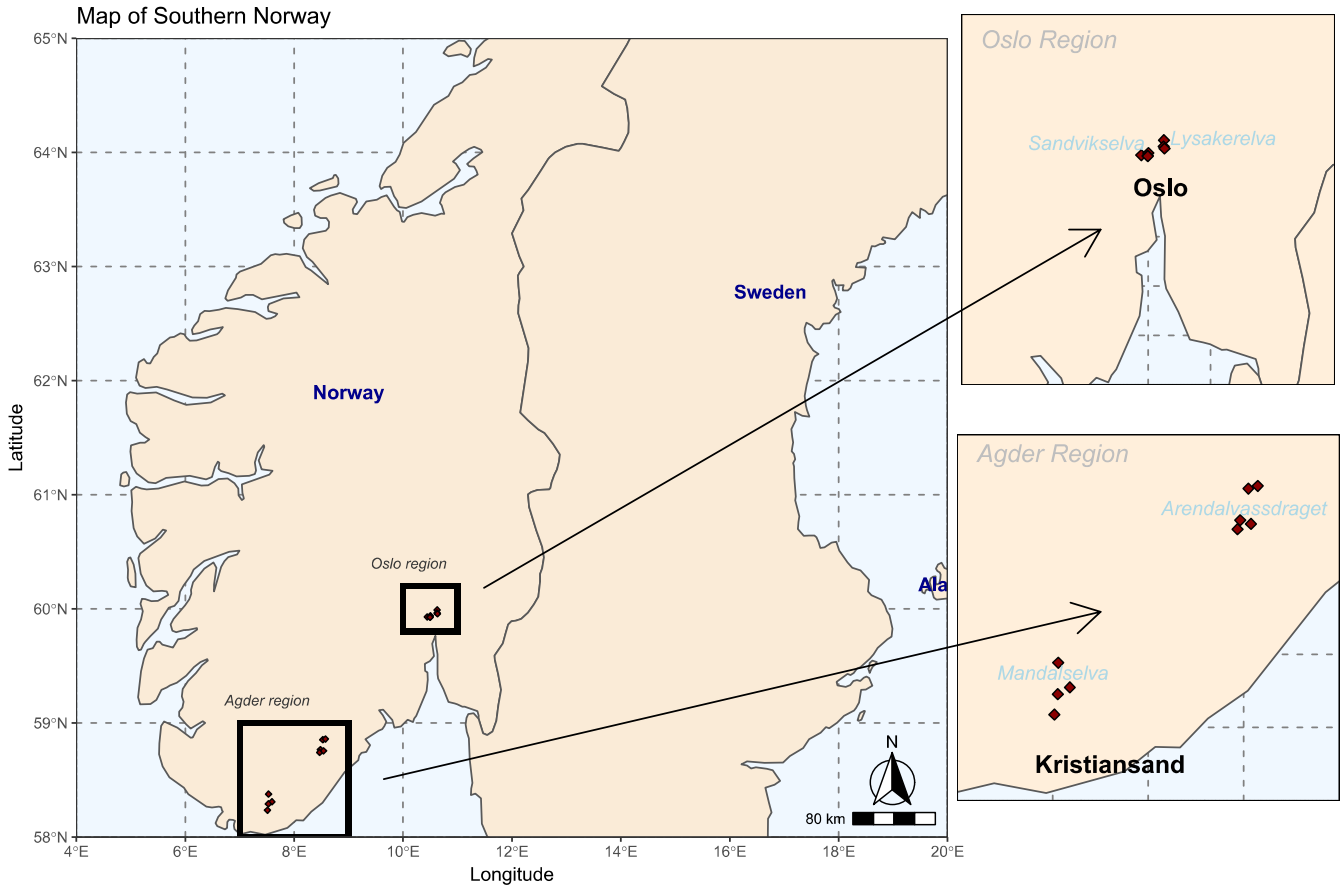


FIGURE 1 Map of Norway and location of the sampling sites in the two regions. Highlighted in the small pictures are the four catchments

biofilm solution was then vortexed and sonicated (20 W for 1 min; Microson XL2000 ultrasonic liquid processor with 1.6-mm-diameter microtip probe, Misonix) to detach the cells from the organic matter. The resulting slurry was left overnight at 4°C, which allowed coarse particles to settle. Thereafter, 1 ml of the resulting slurry was transferred to a 2-ml Eppendorf tube, and 1 ml of the density gradient medium Nycodenz (Nycomed) was placed underneath using a syringe needle. High-speed centrifugation was performed in a swing-out rotor for 90 min at 4°C. Nycodenz-purified subsamples (375 µl) were filtered on 0.2-µm polycarbonate membranes (47 mm diameter, Nuclepore) by gentle vacuum (< 0.2 bar) and washed with 10–20 ml of sterile ultrapure water. One section of each filter was stained for 10 min with DAPI (1 µg ml^{-1} final concentration) and then fixed to a glass slide to be analysed by epifluorescence microscopy. The remaining filter was stored at −20°C for further CARD-FISH analysis.

To quantify the community composition, CARD-FISH was used. The relative abundances for the domains of Bacteria and Archaea, four subphyla of the Proteobacteria (Alpha-, Beta-, Gamma- and Delta-Proteobacteria) and the phylum Firmicutes were obtained.

In situ hybridization was carried out following the protocol of Fazi et al. (2007), Fazi et al. (2013).

Specific oligonucleotidic probes (Biomers), labelled with rRNA-target horseradish peroxidase (HRP), were used to target Bacteria (EUB338 I-III), Archaea (ARCH915), Alphaproteobacteria (ALF968), Betaproteobacteria (BET42a), Gammaproteobacteria (GAM42a), Deltaproteobacteria (DEL495 a-b-c) and Firmicutes (LGC 354a). BET42a and GAM42a served as competitors for each other; for further details on probes see probeBase (Greuter et al., 2016). In addition, the abundance of photosynthetic picoplankton cells (Cyanobacteria) was estimated by their autofluorescence signal as described in Tassi et al. (2018).

The stained filter sections were observed on a Leica DM LB30 epifluorescence microscope (Leica DM LB 30, at 1,000× magnification). At least 300 cells were counted in 10 microscopic fields randomly selected across the filter sections. The relative abundance of hybridized cells was estimated as the ratio of hybridized cells to total DAPI-stained cells.

2.5 | DNA extraction, library preparation and sequencing

For DNA extraction, 15 ml of slurry containing biofilm scraped from each site was homogenized and a subsample of ~0.4 g on average was weighed for each of the sites and then extracted by using the PowerSoil DNA Isolation Kit (Qiagen) by following the manufacturer's instructions. Quality control of the extracted DNA

($1.6 < A_{260} = 280 < 1.8$ and $A_{260} = 230 > 2$) was performed by using a Nanodrop 3300 (Thermo Scientific). The DNA was stored at −20°C in small aliquots (~ 50 µl) until it was sent to DNASense ApS (Denmark) for sequencing.

Sequencing libraries for the V4 region of the 16S rRNA for Archaea and Bacteria were prepared by a custom protocol based on an Illumina protocol (Illumina, 2015).

Up to 10 ng of extracted DNA was used as template for PCR amplification of the Archaea and Bacteria 16S rRNA gene region V4 amplicons. Each PCR (25 µl) contained dNTPs (100 µM of each), $MgSO_4$ (1.5 mM), Platinum Taq DNA polymerase HF (0.5 U per reaction), Platinum High Fidelity buffer (1×) (Thermo Fisher Scientific) and tailed primer mix (400 nM of each forward and reverse primer).

PCR was conducted with the following programme: initial denaturation at 95°C for 2 min, 30 cycles of amplification (95°C for 15 s, 55°C for 15 s, 72°C for 50 s) and a final elongation at 72°C for 5 min.

Duplicate PCRs were performed for each sample and the duplicates were pooled after PCR. The forward and reverse tailed primers were designed according to Illumina (2015) and contain primers targeting the Archaea and Bacteria 16S rRNA gene region V4: [515FB] GTGYCAGCMGCCGCGGTAA and [806RB] GGACTACNVGGGTWTCTAAT (Apprill et al., 2015).

The primer tails enable attachment of Illumina Nextera Indices necessary for sequencing in a subsequent PCR. The resulting amplicon libraries were purified using the standard protocol for Agencourt Ampure XP Beads (Beckman Coulter) with a bead to sample ratio of 4:5. DNA was eluted in 25 µl of nuclease-free water (Qiagen). DNA concentration was measured using a Qubit dsDNA HS Assay kit (Thermo Fisher Scientific). Gel electrophoresis using Tapestation 2200 and D1000/High sensitivity D1000 screentapes (Agilent) was used to validate product size and purity of a subset of sequencing libraries. Sequencing libraries were prepared from the purified amplicon libraries using a second PCR. Each PCR (25 µl) contained PCRBIO HiFi buffer (1×), PCRBIO HiFi Polymerase (1 U per reaction) (PCRBiosystems), adaptor mix (400 nM of each forward and reverse) and up to 10 ng of amplicon library template. PCR was conducted with the following programme: initial denaturation at 95°C for 2 min, eight cycles of amplification (95°C for 20 s, 55°C for 30 s, 72°C for 60 s) and a final elongation at 72°C for 5 min. The resulting sequencing libraries were purified using the standard protocol for Agencourt Ampure XP Beads (Beckman Coulter) with a bead to sample ratio of 4:5. DNA was eluted in 25 µl of nuclease-free water (Qiagen).

DNA concentration was measured using aQubit dsDNA HS Assay kit (Thermo Fisher Scientific). Gel electrophoresis using Tapestation 2200 and D1000/High sensitivity D1000 screentapes (Agilent) was used to validate product size and purity of a subset of sequencing libraries.

The purified sequencing libraries were pooled in equimolar concentrations and diluted to 2 nM. The samples were paired-end sequenced (2 × 300 bp) on a MiSeq (Illumina) using a MiSeq Reagent kit version 3 (Illumina) following the standard guidelines for preparing and loading samples on the MiSeq. A > 10% PhiX control library was spiked to overcome low complexity issues often observed with amplicon samples.

The files received from the sequencing agency had already been demultiplexed, so the tags (FWD and REV) identifying each sample were absent from the sequences and did not have to be removed during the first filtering step. The fastq files (16 R1 and 16 R2) were checked for quality by using FASTQC software (version 0.11.4; Andrews, 2010) to inspect the overall quality of the sequences and look for primers, adapters and Ns content.

The inspection revealed no presence of Ns in the sequences, no presence of adapters and good overall quality of the sequences. By looking at the overrepresented sequences, we found out that the FWD primer was in the R2 files, while the REV primer in the R1.

The demultiplexed sequences were processed in r 3.5.1 (R Development Core Team, 2008) by first using CUTADAPT 1.14 (Martin, 2011) to trim the primers from the sequences.

The primers were identified by creating two objects, one for the FWD and one for the REV primer. Subsequently a function was created to detect all possible orientations for the primers.

Next, a function was applied to check the number of times the primers appeared in the forward and reverse read, while considering all possible primer orientations. Finally, the FWD, REV and their complements were trimmed off the sequences by using CUTADAPT. To ensure a good outcome of the trimming step, the primer count was run again on the sequences processed with CUTADAPT and no primers were found in all possible sequence orientations. Once the primers were trimmed, we used DADA2 (version 1.10.1) (Callahan et al., 2016) to construct amplicon sequence variants (ASVs).

Taxonomic assignment to the ASVs was made by using the "assignTaxonomy" function, which is based upon the naive Bayesian classifier method (Wang et al., 2007). The input for this command is the set of ASVs to be classified and a training set of reference sequences with known taxonomy; we used the "silva_nr_v132_train_set.fa" (Callahan, 2018).

After taxonomic assignment, we ran the "assignSpecies" command to assign species-level taxonomy with more accuracy by using the "silva_species_assignment_v132.fa" database (Callahan, 2018). As stated in Edgar (2017), the only proper threshold for species-level taxonomic assignment to HTS 16S amplicon data is 100% identity for ASVs.

A total of 884,353 reads were obtained for the 16 sampling sites from the Illumina sequencing platform after the "pre-filtering" step where primers and ambiguous bases were removed. Each sample had on average 55,272 reads, with a minimum of 30,324 at AML4 and a maximum of 66,571 at HAUG. The final number of total sequences, after being checked for quality and chimeric sequence removal, was 617,816. On average, the samples after bioinformatic processing had 38,613 reads, 11,650 was the minimum number still at AML4 while the maximum final value of 48,196 reads was for BOG2 (Table S2).

The final number of reads after taxonomic assignment and after removal of sequences belonging to Eukaryota and Chloroplast ranged from 47,458 at SAND to 10,130 at AML4.

Raw sequences are deposited at the European Nucleotide Archive under accession nos. ERR4650589 to ERR4650605.

2.6 | Statistical analysis

Statistical analyses were performed in r, version 3.5.1 (R Development Core Team, 2008).

The environmental parameters were tested for normality using Shapiro-Wilks; only NH_4 was log-transformed ($logNH_4$) to meet normality. We tested for multicollinearity using the correlation matrix and computed the variance inflation factor (VIF) and the tolerance statistic. The analysis led us to select a few noncollinear parameters: TOC, TN, TP, pH, $logNH_4$. To visualize the principal environmental gradients, a principal component analysis (PCA) was performed. The environmental variables were standardized before analysis. To test the difference between the two regions analysed we performed *t* tests for the selected environmental parameters.

To analyse the prokaryotic community structure three tables were created, one with ASV abundances and, from the CARD-FISH results, one with absolute abundances and one with relative abundances. To achieve equal sampling depth, we rarefied (randomly subsampling) the ASVs to the same library size number (n = 10,130, minimum number of total sequences found).

From the rarefied and standardized (by using the "decostand" function and Hellinger method) ASV abundances a Bray–Curtis dissimilarity matrix was created using the "vegdist" function. For visualization of the prokaryotic community distribution, nonmetric multidimensional scaling ordination (NMDS) was performed using the Bray–Curtis dissimilarity matrix. Starting from an initial configuration we produced 100 configurations, using the "global" model (Liu et al., 2008), and 200 as the maximum number of iterations. Unreliable distances (B-C > 0.9) were replaced by geodesic distances using a step-across method to calculate the shortest distance on any kind of "underlying nonlinear structure" (Williamson, 1978).

We extracted the two best solutions, those with the lowest stress value, and then scaled the axes of both the solutions to half change units and varimax rotation by using the "postmds" function. To assess the fit between the two best NMDS found, we used the Procrustes comparison analysis and the "protest" function. The protest statistics (Sums of Square Difference [SSD] =1.144e-11; r = 1; permutation test [999] =0.001) confirmed the fit between the two best NMDS found. We then used the "envfit" function to fit the environmental parameters, used to produce the PCA, to see which variable was driving the community composition of microbes most. The ordination diagram was then built with the best solution overall, with the fitted values for the water physicochemical parameters. To test for differences in the microbial community structure, we performed an analysis of similarities (ANOSIM) using the catchments as the factorial variable. This type of analysis provides statistical information on the difference between microbial communities according to the grouping variable.

We also performed a Beta dispersion analysis to test if the prokaryotic abundances from CARD-FISH and 16S sequencing were homogeneously dispersed among groups of two different factorial variables: rivers and regions. By using the "adonis" function from the VEGAN package (Oksanen et al., 2013), we also tested the species compositional difference between the factorial variables.

Mantel tests were run between Bray–Curtis similarity matrices for ASV abundances, CARD-FISH relative and absolute abundances, and the Euclidean distance matrix for the standardized environmental variables to detect similar patterns and thus the driving variables for the bacterial community composition.

From the ASV abundance tables at class and genus levels, we calculated the prevalence (occurrence for sampling site) and coefficient of variation (standard deviation of taxon abundance divided by the mean) to detect taxa which might be suitable as biological indicators. These two parameters were plotted against each other in a scatter plot to find taxa with the highest prevalence and highest variation. Taxa with high prevalence and variance could be used as indicators of environmental gradients. We performed a redundancy analysis (RDA) with the ASV abundance matrix at class and genus levels and the environmental variables used for the PCA to detect any relationship among the taxa with highest prevalence and variance and the environmental variables. The class-level ASV abundance matrix was transformed using the Hellinger method to reduce the effect of large abundances.

3 | RESULTS

3.1 | Regional vs catchment characteristics

3.1.1 | Water physicochemical characteristics

The water chemistry parameters for the two regions are shown in Table S3. The two regions showed different patterns (Figure 2), but the main difference was pH, which in the Oslo region was on average 7.53 ± 0.42, while the mean value for Agder was 5.70 ± 0.40.

Mean conductivity, measured in the rivers from the Oslo region, was 86.86 ± 37.77 μS cm^{-1}, considerably higher than the mean values recorded in Agder (19.20 ± 9.57 μS cm^{-1}). Another marked difference between the two regions analysed was the mean value for sulphate, showing higher concentration in the rivers from Oslo (5.7 ± 1.73 mg L^{-1} on average). The average value for sulphate in Agder was 1.32 ± 0.70 mg L^{-1}. Among the parameters showing variability between the rivers from southeastern and southwestern Norway, one of the most important was TOC, reaching a mean value of 7.56 ± 1.32 mg L^{-1} in the Oslo region (southeast), but only 5.42 ± 1.15 mg L^{-1} in Agder (southwest).

Measurements for TN and TP showed a similar pattern, with higher concentrations in the catchments from the Oslo region (on average 574.3 ± 20.7 μg L^{-1} TN, 9 ± 2.7 μg L^{-1} TP). The average concentration for TN for Agder was 345.6 ± 181.6 μg L^{-1}. The average TP concentration was 5.4 ± 3.1 μg l^{-1} for Oslo and 3.6 μg L^{-1} for

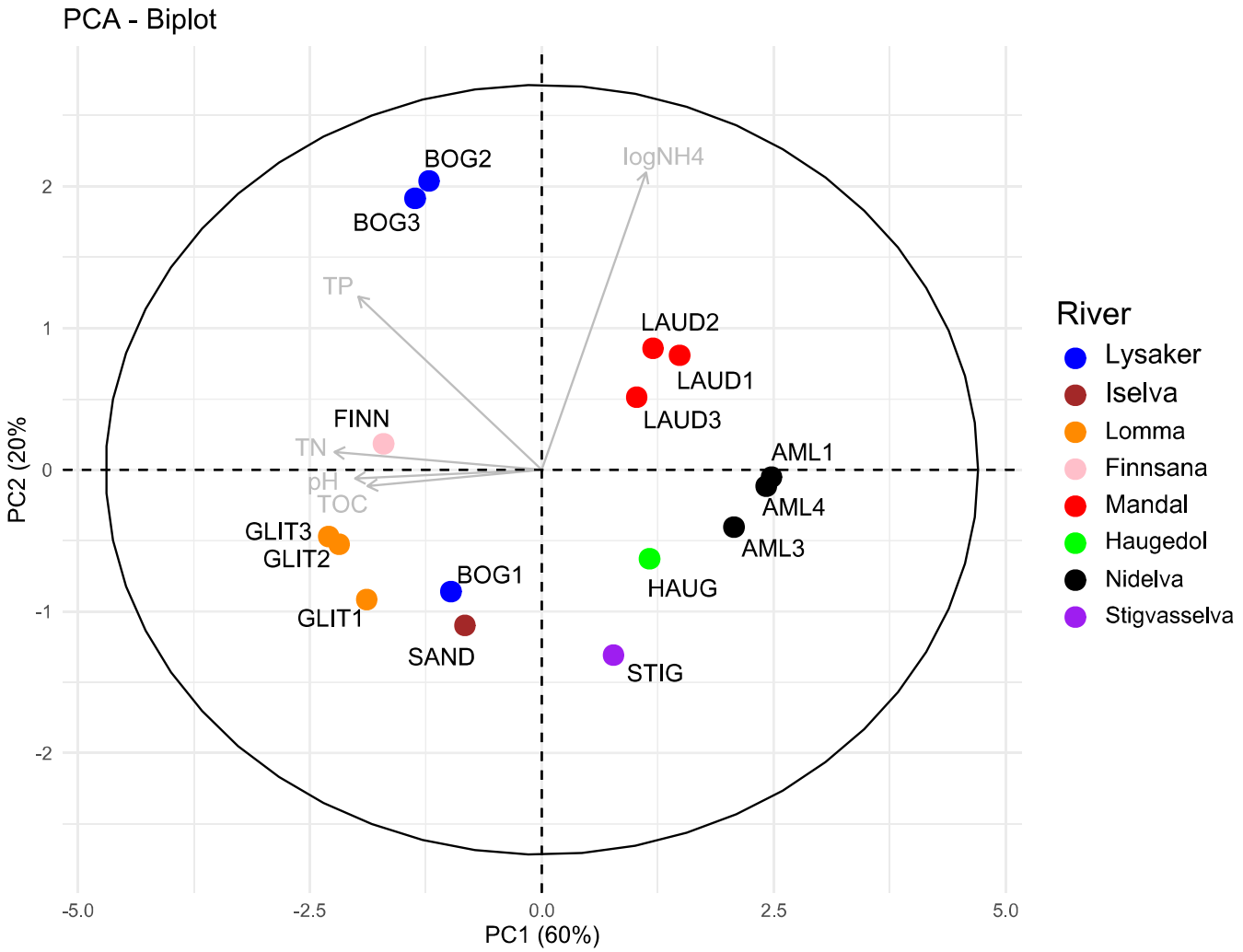


FIGURE 2 PCA with the relevant standardized environmental variables. Length of vectors is proportional to the contribution of the variable to the principal components. Colours correspond to the sampled rivers

Agder. As confirmed by the *t* test result ($p < .001$), pH was the variable demarcating the two regions. Results for *t* tests were significant also for TN ($p < .05$), TP ($p < .05$) and TOC ($p < .05$) but not for $logNH_4$ ($p = .3$).

3.1.2 | Archaeal and bacterial abundances

Total prokaryotic abundance, as determined by DAPI-stained cell counts, ranged from a maximum of $3.99 \times 10^{10} \pm 1.23 \times 10^{9}$ cells g^{-1} at SAND, to a minimum of $2.71 \times 10^{9} \pm 3.86 \times 10^{8}$ cells g^{-1} at BOG1. Among all DAPI-stained cells we could affiliate on average 84.9 ± 3.76% to Bacteria and 4.7 ± 0.86% to Archaea. The highest abundances for both Bacteria ($3.60 \times 10^{10} \pm 7.36 \times 10^{8}$ cells g^{-1}) and Archaea ($1.76 \times 10^{9} \pm 4.35 \times 10^{8}$ cells g^{-1}) were found at SAND, whereas the lowest abundance (respectively $2.38 \times 10^{9} \pm 5.48 \times 10^{7}$ cells g^{-1} for Bacteria and $1.39 \times 10^{8} \pm 5.24 \times 10^{7}$ cells g^{-1} for Archaea) were found at BOG1. Further detail are given in Table S4.

Overall, the highest abundances for Alphaproteobacteria ($1.37 \times 10^{10} \pm 1.67 \times 10^{9}$ cells g^{-1}), Beta- ($1.47 \times 10^{10} \pm 1.62 \times 10^{9}$ cells g^{-1}) and Gammaproteobacteria ($1.58 \times 10^{9} \pm 3.66 \times 10^{8}$ cells g^{-1}) followed the same pattern as the total prokaryotes abundances, being highest at SAND. The lowest values, for Alphaproteobacteria ($6.48 \times 10^{8} \pm 5.78 \times 10^{7}$ cells g^{-1}), Betaproteobacteria ($7.08 \times 10^{8} \pm 1.10 \times 10^{8}$ cells g^{-1}) and for Gammaproteobacteria ($2.29 \times 10^{8} \pm 3.26 \times 10^{7}$ cells g^{-1}) were recorded at BOG1. The highest abundances for the Deltaproteobacteria were found at FINN ($3.06 \times 10^{9} \pm 8.36 \times 10^{8}$ cells g^{-1}), while the lowest ($3.21 \times 10^{8} \pm 1.49 \times 10^{7}$ cells g^{-1}) were recorded at BOG1. Firmicutes showed the highest cell abundances at LAUD1 ($8.77 \times 10^{8} \pm 3.09 \times 10^{8}$) and the lowest at BOG1 ($7.76 \times 10^{7} \pm 1.11 \times 10^{7}$ cells g^{-1}). Autofluorescence was highest at BOG3 ($1.56 \times 10^{9} \pm 2.65 \times 10^{8}$ cells g^{-1}), and lowest at STIG ($2.77 \times 10^{8} \pm 1.14 \times 10^{8}$ cells g^{-1}) (Figure 3). Detailed information on the abundances of the specific bacterial groups analysed is presented in Table S5. By using the data from only those rivers impacted by hydropower and dams, we

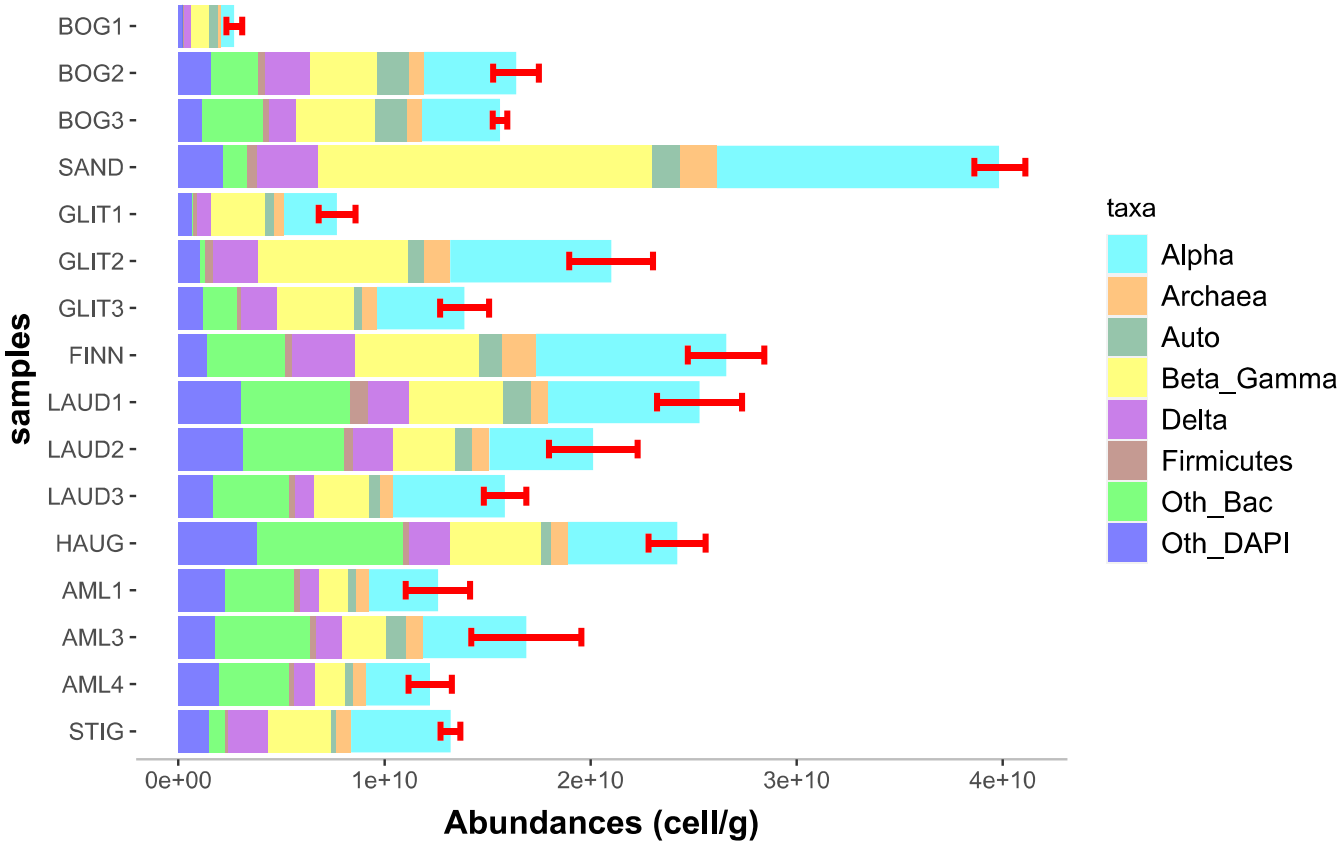


FIGURE 3 Absolute cell numbers for the prokaryotic taxa analysed by CARD-FISH. The key shows the classes Proteobacteria (Alpha-, Beta- and Gamma-[joint abundances] and Delta-); the phylum Firmicutes (Firmicutes); the autofluorescent cells (Auto), which correspond to photosynthetic prokaryotes; archaeal abundances and the proportion of Bacteria not identified by our probes (Oth_Bac); Oth_DAPI refers to the DAPI-stained cells which were not identified by either the bacterial or the archaeal probes

plotted the abundances for all bacterial groups (Figure 4), and shows a pattern between the ratios of Alphaproteobacteria and Gammaprotebacteria. These bacterial classes were dominant in the rivers from southeast Norway, both having similar cell abundances. The rivers from southwest Norway showed a different pattern, being mainly dominated by Alphaproteobacteria in terms of abundance, while the Gammaproteobacteria and the other classes were considerably less abundant compared to in rivers from the Oslo region.

3.1.3 | Bacteria diversity

Bacterial community composition showed that sequences were affiliated with 25 bacterial phyla and two archaeal phyla (Thaumarcheota and Euryarcheota). Overall the dominant phylum was the Proteobacteria, whoch accounted for 62.1% of the sequences on average. Within the phylum Proteobacteria, Gammaproteobacteria represented the most abundant class on average (43.6%) with Alphaproteobacteria representing the second most abundant class with 17.8% on average. Other abundant bacterial phyla were the Firmicutes (on average 9.2%), among which the most abundant class was the Bacilli with 8.5% on average, Cyanobacteria (7.4%) and Bacteroidetes (5.6%).

The most abundant genera belonged to the class Gammaproteobacteria: *Pseudomonas* (9.4% on average), *Acinetobacter* (8.3%), *Yersinia* (6.7%) and *Lactococcus* (6%).

The NMDS plot shows a clear clustering of the sampling sites according to geographical distribution, dividing the samples from the region around Oslo from those from Agder (Figure 5). This is consistent with the cluster analysis and the PCA conducted on the environmental variables. The Mantel test (Table 1) showed a significantly positive correlation between dissimilarity matrices of the ASV community composition (HTS) and environmental parameters (r = .687, p < .001). Among the single environmental parameters, the one showing the highest significant correlation was TN (r = .755, p < .001). A positive correlation was derived between the dissimilarity matrices of community composition obtained from ASVs and the relative abundances of CARD-FISH (CARD PERCENTAGES)-targeted groups (r = .4, p < .05) while the correlation with the absolute abundances of FISH (CARD BAB)-targeted groups was weak (r =.184, p <.05).

To visualize the underlying trends in our ASV abundances, we performed an "envfit" analysis to fit the environmental features, highlighted in the PCA for water chemistry, to the NMDS created

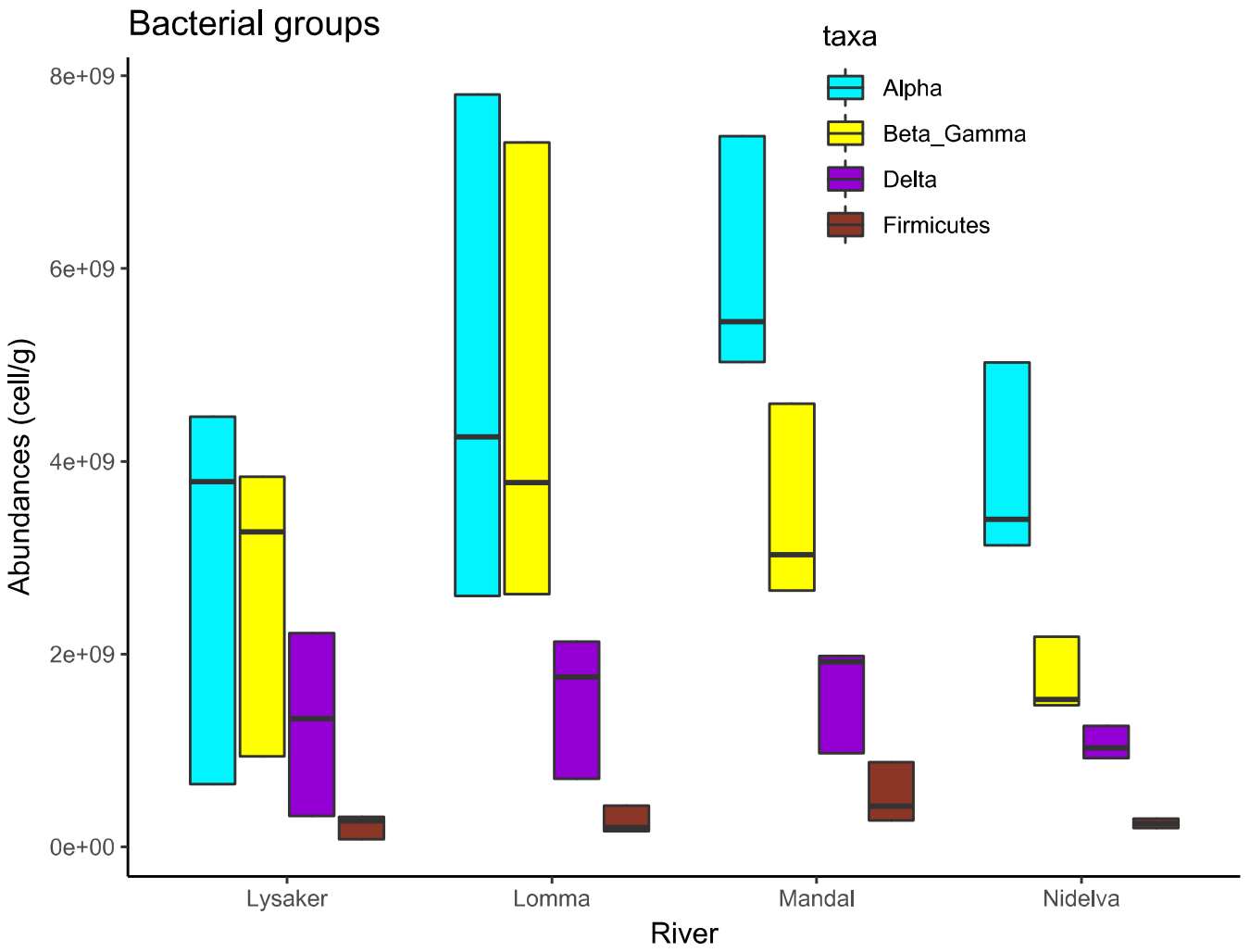


FIGURE 4 Comparison of cell numbers for the heterotrophic bacterial taxa analysed by CARD-FISH in the four rivers impacted by dams and hydropower plants (Alphaproteobacteria, Beta- and Gammaproteobacteria [joint abundances], Deltaproteobacteria and Firmicutes). Different ratios between Alpha- and Beta_Gammaproteobacteria groups were detected for the Oslo and the Agder regions

from the sequencing data. The correlation results confirmed that community composition and pH were closely associated (r =.9, p <.001) (Table S6).

3.1.4 | Comparison among methods and bacterial indicators

The results of the ANOSIM performed on the Bray–Curts matrix obtained from ASVs showed a significant association with specific catchment (r =.8, p <.001). This was confirmed by Beta Dispersion Analysis where we detected variation in the species composition at both, regional (PERMANOVA, r^2 =.274, p <.001) and catchment scale (PERMANOVA, r^2 =.461; p <.001). In comparison, the CARD-FISH results showed very high variability, and thus no significant association with specific regions or catchments could be observed.

Analysis of prevalence and the coefficient of variation at the class level showed that some taxa (at class and genus level) were distributed across all sampling sites (Figure 6a,b; Figure S1a,b), showing large variability in their abundances. The highly variable taxa with wide prevalence across the catchments and two geographical regions showed distinct patterns as revealed by RDA (Figure 6c). For example, Bacilli was positively associated with TP and Bacteroidia with ammonium, as well as Alphaproteobacteria with low pH and Gammaproteobacteria with high values for TN and TOC. Similarly, using abundance data at the the genus level revealed associations of *Janthiniobacterium* with TP and *Sphingomonas* with low values of TN and TP (Figure S1c). As such these taxa may provide biological indicators for the status of the river system with regard to nutrients and acidification.

4 | DISCUSSION

This study casts new light into the prokaryotic community structure of epilithic biofilms dwelling in rivers affected by natural and anthropogenic impacts. The combination of two techniques,

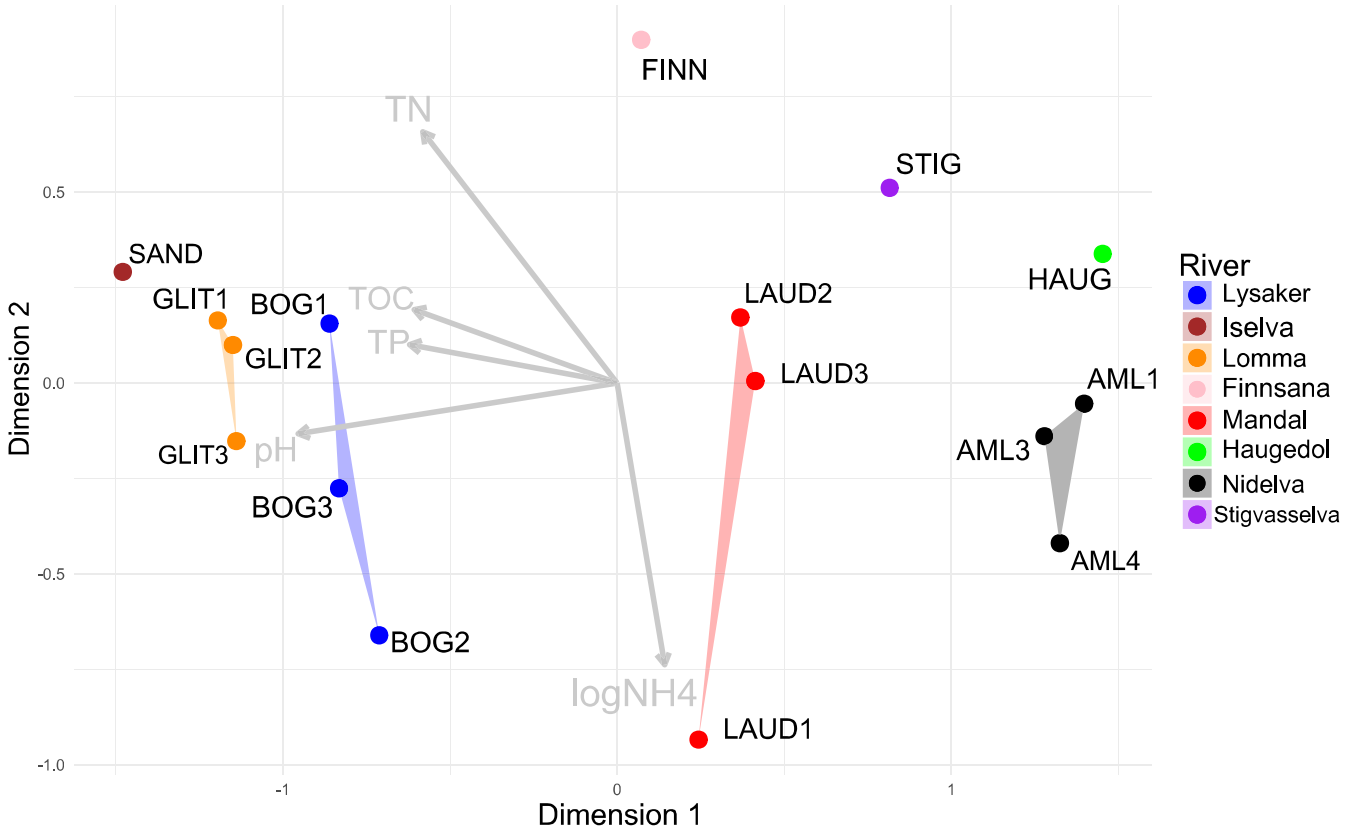


FIGURE 5 To analyse the distribution of the biofilm community structure and its relationship with the environmental parameters we plotted the envfit analysis produced by using the best GNMDS out of 100 iterations, performed on the Bray–Curtis matrix of ASV abundances and the data frame for the variables used in the PCA. (Stress value =0.06)

16S rRNA sequencing and CARD-FISH, allowed us to gain insight into the community composition of epilithic prokaryotes, which are still poorly understood in the context of riverine ecosystems. Quantitative results such as those obtained by CARD-FISH allow us to detect variations in actual cells numbers and activity of specific taxa, which would be otherwise lost by analysing only the number of gene copies provided by sequencing (Fazi et al., 2020). The CARD-FISH results obtained in our study revealed a great variation at small spatial scales, such as in biofilms belonging to the same river, and thus seems to detect in-system variability in the microbial community composition to a greater extent than metabarcoding. However, this extreme microscale variability might mask the overall effects of the main drivers for the whole microbial community composition. With regard to analysing communities occurring across large spatial scales, the sequencing methods used in our study have proven their validity. Metabarcoding provides a huge amount of data with high taxonomic resolution, which can be related to the physicochemical parameters of the environment (Ligi et al., 2014). It enables the exploration of large-scale patterns in relation to environmental conditions and a finer taxonomic resolution than hybridization methods (Bouvier & del Giorgio, 2003; Corte et al., 2013).

Both techniques revealed a dominance of Proteobacteria across all the samples corroborating most previous studies on freshwater epilithic bacterial communities (Battin et al., 2016; Besemer et al., 2012; Wilhelm et al., 2013). Overall, the total prokaryotic cell abundances obtained by DAPI staining were an order of magnitude lower than those found by Fazi et al. (2005), but comparable to those found by Zoppini et al. (2010) in similar freshwater systems. Beta- and Alphaproteobacteria were the most abundant classes according to the CARD-FISH results, in line with the results of studies on microbial communities in urban streams (Araya et al., 2003) and freshwater mesocosms (Lupini et al., 2011). Gamma- and Deltaproteobacteria were less abundant, similar to the findings of Webster et al. (2004), where biofilms at different stages of development were analysed by FISH (fluorescence *in situ* hybridization). From the absolute abundances of bacterial groups obtained by CARD-FISH, we were also able to detect peaks of bacterial cell numbers, which may be related to pollution sources and impacts that would not have been identified by using sequencing alone (Freixa et al., 2016; Bakenhus et al., 2019). Nevertheless, the variability between sampling sites, within the same river, was too large to identify any associated variables providing potential explanations. We speculate that cell numbers might be affected by microscale ecological features that we did not measure, such as interactions with other biological or physicochemical components varying at the microscale. Variability might also be caused by random events such technical or sampling biases.

TABLE 1 Mantel test covariance coefficients based on Spearman rank correlation between the Bray–Curtis matrix for ASV abundances and Euclidean matrices for the environmental variables. Correlation between the Bray–Curtis similarity matrices for CARD-FISH percentages and abundances (CARD BAB) with metabarcoding and environmental parameters

Mantel test		
Permutations =9,999		
Mantel statistic based on Spearman's rank correlation rho		
Metabarcoding vs. single environmental variable	**Mantel test *r***	***p***
TN	.755	1e−04***
ph	.723	1e−04***
TP	.332	.0064***
TOC	.321	.0076***
$LogNH_4$	.138	.0745
Metabarcoding vs. standardized environmental variables matrix	.687	1e−04***
CARD percentage vs. standardized environmental variables matrix	.407	.001***
CARD BAB vs. standardized environmental variables matrix	.2887	.015***
Biological matrices		
Metabarcoding vs. card percentages	0.389	.002***
Metabarcoding vs. card bab	0.184	.038***

***p <.001, **p <.05.

The metabarcoding analysis revealed that among the Proteobacteria, the most well-represented group was the class Gammaproteobacteria based on abundances. This result might seem to contradict the results from *in situ* hybridization, but according to the taxonomic database "silva_nr_v132_train_set.fa" (Callahan, 2018), the class Betaproteobacteria, formerly part of the class Gammaproteobacteria, is now an order. This new classification might explain the discrepancy with regard to Gammaproteobacteria between the two methods and reason why we joint the abundances of Betaproteobacteria and Gammaproteobacteria for the visualization of the CARD-FISH results. The most well-represented genera belonging to the class Gammaproteobacteria in the epilithic biofilms were *Pseudomonas*, *Acinetobacter* and *Yersinia*. *Pseudomonas* and *Acinetobacter* are well-known members in the early stages of biofilm successions in marine environments (Lee at al., 2008), and the latter is also common in soils and freshwaters (Williams et al., 1996). In terms of read numbers, the second most important bacterial class was the Alphaproteobacteria with *Sphingomonas* being the most common genus. *Sphingomonas* has been previously found to be an important player in biofilm structural composition because of the high production of expolysaccharides, a major constituent of microbial biofilms (Johnsen et al., 2000).

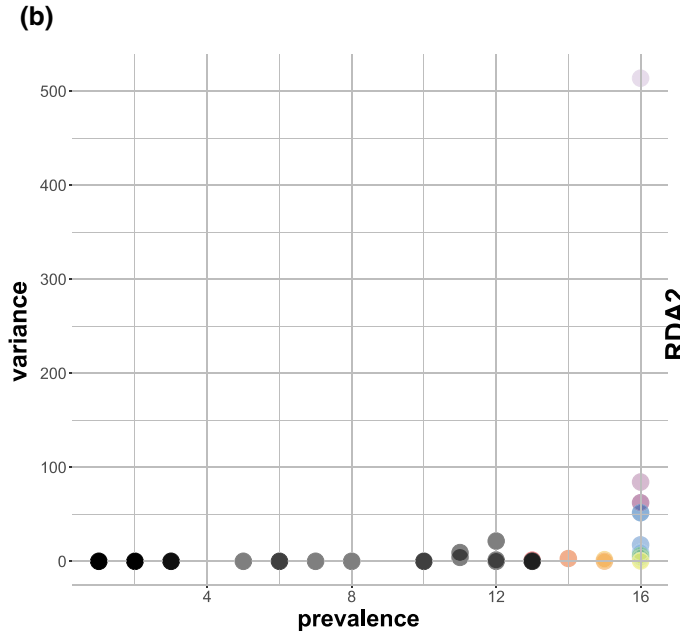
Betaproteobacteria were the third proteobacterial group to be highly represented in the sequencing data, with *Massilia* and *Janthiniobacterium* being the most common genera for this class. Betaproteobacteria are the group most associated with freshwater ecosystems, including important functional groups of bacteria such as ammonia oxidizers, which are vital in the global nitrogen cycle (Barberán & Casamayor, 2010; Sekar et al., 2004; Zhang et al., 2012). *Massilia* and *Janthiniobacterium*, both belonging to the order Burkholderiales, are typical of freshwater environments (Gołębiewski et al., 2017). *Massilia* is a ubiquitous genus, often present in soils and in biofilms and exhibiting unique properties including expolysaccharide production, incredible adhesive force and hydrophobicity, making biofilms more resistant (Liu et al., 2012).

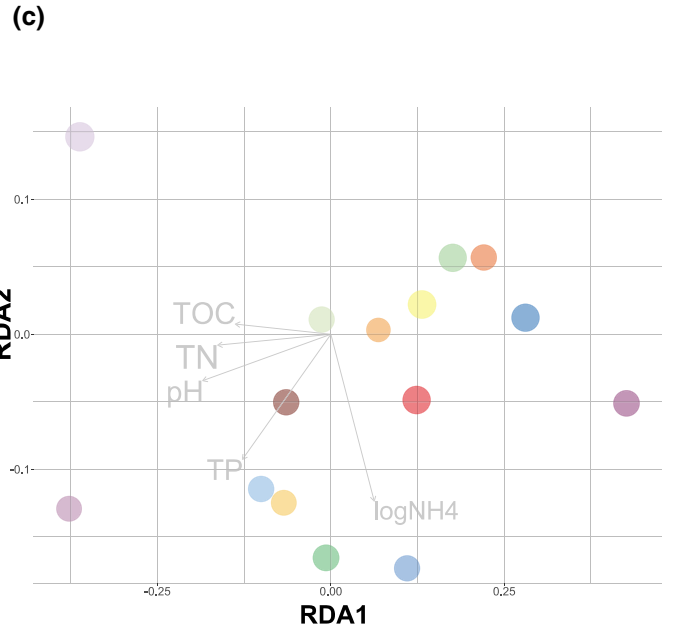
According to our findings, the microbial community structure is profoundly dependent on the physicochemical features of the region, confirming previous results on microbial communities from sediments in coastal areas, estuaries and rivers (Freixa et al., 2016; Aylagas et al., 2017; Borja, 2018; Fazi et al., 2020). While several environmental characteristics were associated with the epilithic community dynamics, the driving environmental parameter appears to be the acidity of water (as confirmed by the pH results of the envfit analysis, $r^2 = 0.9$, p <.001), which is considerably lower in the southwestern region of Norway. It is well known that pH can influence microbial communities favouring certain strains such as members of the Alphaproteobacteria (Bragina et al., 2012; Dedysh, 2009; Goffredi et al., 2011), which were dominant in the region of Agder, where rivers had on average lower pH. In addition to the more acidic environment, the nutrient load in the rivers from Agder was generally much lower compared to the rivers flowing through the Oslo area. This is due to different anthropogenic pressures in the two regions (Peder Flaten, 1991; Nordeidet et al., 2004; Reimann et al., 2009; Johannessen et al., 2015). This characteristic might also have affected the ratio between Alpha- and Gammaproteobacteria, which in the Oslo region displayed similar cell numbers, whereas in Agder the ratio was consistently different. Overall, study regions confounded the relationships between microbial community structure and environmental variables because of their distinct differences in water chemistry. So, while our study was able to show a strong response to a number of environmental variables, we are not able to disentangle this from regional effects which would need inclusion of more regions and more sampling sites.

Our study had some limitations, and true replicates are needed to get indicator taxa. However, the comparison between the two

FIGURE 6 (a) Best represented prokaryotic classes among all the sampling sites. (b) Prokaryotic classes with highest prevalence among the sampling sites and highest variance; the 15 coloured dots are the taxa most suitable as biological indicators (highlighted in the ellipses) given their broad presence and wide variance across different environments. (c) RDA showing the distribution patterns for the 15 identified taxa and the environmental parameters


(a)
BOG1
BOG2
BOG3
SAND
GLIT1
GLIT2
GLIT3
FINN
LAUD1
LAUD2
LAUD3
HAUG
AML1
AML3
AML4
STIG
samples
0
25
50
75
100
proportion of reads [%]
Class
Gammaproteobacteria
Bacilli
Oxyphotobacteria
Alphaproteobacteria
Bacteroidia
Actinobacteria
Blastocatellia_(Subgroup_4)
Verrucomicrobiae
Planctomycetacia
Deltaproteobacteria
Deinococci
Phycisphaerae
Acidobacteriia
Anaerolineae
Gemmatimonadetes


(b)
500
400
300
200
100
0
variance
4
8
12
16
prevalence


(c)
0.1
0.0
-0.1
RDA2
TOC
TN
pH
TP
logNH4
-0.25
0.00
0.25
RDA1

techniques, besides showing corresponding patterns, provide different insights into the complexity of the prokaryotic community structure of riverine epilithic biofilms. Sequencing allowed us to detect the deep diversity among the microbial taxa dwelling in different river systems, with higher taxonomic resolution than with CARD-FISH. CARD-FISH provided absolute cell numbers for specific prokaryotic groups, which is the only quantitative way, based on absolute cell numbers, to assess the composition of microbial communities (Bakenhus et al., 2019; Corte et al., 2013). CARD-FISH showed high variability at the microscale, highlighting patterns between the bacterial groups analysed that were not evident from the metabarcoding results.

Overall, our results suggest that sequencing is better suited than CARD-FISH to assess overall community dynamics. On the other hand, hybridization *in situ* is extremely valuable in later stage studies, aiming to analyse target taxa (i.e., indicators for pollution, diseases, eutrophication, etc.). Consequently, the use of a specific technique parallels the experiences gained from other biological groups, such as macroinvertebrates, where methods, including taxonomic resolution and enumeration, differ depending on the type of bio-assessment or scientific aims (Friberg, 2014).

4.1 | Future perspectives

Here, we show how new microbial indicators can be provided by looking at the ratios between coefficients of variation and prevalence of prokaryotic taxa detected by 16S rRNA sequencing and by using absolute abundances from CARD-FISH as a conversion factor to correct for the relative read abundances (Figure S2). By associating ASVs at specific taxonomic levels with environmental properties we might also be able to detect prokaryotic biological indicators to be used in setting environmental quality thresholds in aquatic and terrestrial environments. Our results also indicated that communities could differ substantially between geologically distinct regions, emphasizing the need to use a reference conditions approach (*sensu* Water Framework Directive [WFD]) in future biomonitoring with microbial indicators. While the scope of our study was too limited to establish generalized relationships between environmental variables and microbial indicators, it strongly implied that such relationships indeed exist and could be the backbone of powerful bioindicator tools for the future, filling in the black box that currently exists with regard to large parts of the microbial communities in rivers.

CONFLICT OF INTERESTS

The authors declare that they have no known competing financial interests or personal relationships that influenced the work reported in this paper. Exception: Alexander Eiler is the owner and founder of eDNA solutions AB specialized on bioinformatics and eDNA research and innovation. He was also a part-time employee of eDNA solutions AB parallel with this research project conducted as part of his appointment as a Professor at the University of Oslo.

AUTHOR CONTRIBUTIONS

LP and NF conceived the study and the experimental design. LP worked on the acquisition of data. LP analysed the data. LP, AE, SF and NF discussed the results. LP, AE, SF and NF wrote the manuscript.

DATA AVAILABILITY STATEMENT

CARD-FISH data, environmental variables and factorial variables used for this paper are available at https://osf.io/f4zb2/?view_only=dfbefd13e28140728da421de1b999922. Identifier: https://doi.org/10.17605/OSF. IO/F4ZB2. DNA sequences: European Nucloeotide Archive Accession nos. ERR4650589 to ERR4650601 and ERR4650603 to ERR4650605.

ORCID

Lorenzo Pin https://orcid.org/0000-0003-0964-2424

REFERENCES

Amalfitano, S., & Fazi, S. (2008). Recovery and quantification of bacterial cells associated with streambed sediments. *Journal of Microbiol Methods*, *75*, 237–243.

Andrews, S. (2010). FastQC: a quality control tool for high throughput sequence data. http://www.bioinformatics.bbsrc.ac.uk/projects/fastqc.

Apprill, A., Mcnally, S., Parsons, R., & Weber, L. (2015). Minor revision to V4 region SSU rRNA 806R gene primer greatly increases detection of SAR11 bacterioplankton. *Aquatic Microbial Ecology*, *75*, 129–137.

Araya, R., Tani, K., Takagi, T., Yamaguchi, N., & Nasu, M. (2003). Bacterial activity and community composition in stream water and biofilm from an urban river determined by fluorescent in situ hybridization and DGGE analysis. *FEMS Microbiology Ecology*, *43*(1), 111–119. https://doi.org/10.1111/j.1574-6941.2003.tb01050.x.

Arthington, A. H., Bhaduri, A., Bunn, S. E., Jackson, S. E., Tharme, R. E., Tickner, D., Young, B., Acreman, M., Baker, N., Capon, S., Horne, A. C., Kendy, E., McClain, M. E., Poff, N. L. R., Richter, B. D., & Ward, S. (2018). The Brisbane declaration and global action agenda on environmental flows (2018). *Frontiers in Environmental Science*, *6*, 45.

Aylagas, E., Borja, Á., Tangherlini, M., Dell'Anno, A., Corinaldesi, C., Michell, C. T., Irigoien, X., Danovaro, R., & Rodríguez-Ezpeleta, N. (2017). A bacterial community-based index to assess the ecological status of estuarine and coastal environments. *Marine Pollution Bulletin*, *114*(2), 679–688. https://doi.org/10.1016/j.marpolbul.2016.10.050.

Bakenhus, I., Wemheuer, B., Akyol, P., Giebel, H.-A., Dlugosch, L., Daniel, R., & Simon, M. (2019). Distinct relationships between fluorescence in situ hybridization and 16S rRNA gene- and amplicon-based sequencing data of bacterioplankton lineages. *Systematic and Applied Microbiology*, *42*(5), 126000. https://doi.org/10.1016/j.syapm.2019.06.005.

Baranov, V., Jourdan, J., Pilotto, F., Wagner, R., & Haase, P. (2020). Complex and nonlinear climate-driven changes in freshwater insect communities over 42 years. *Conservation Biology*, https://doi.org/10.1111/cobi.13477.

Barberán, A., & Casamayor, E. (2010). Global phylogenetic community structure and β-diversity patterns in surface bacterioplankton metacommunities. *Aquatic Microbial Ecology*, *59*, 1–10. https://doi.org/10.3354/ame01389.

Battin, T. J., Besemer, K., Bengtsson, M. M., Romani, A. M., & Packmann, A. I. (2016). The ecology and biogeochemistry of stream biofilms. *Nature Reviews Microbiology*, *14*(4), 251–263. https://doi.org/10.1038/nrmicro.2016.15.

Battin, T. J., Kaplan, L. A., Findlay, S., Hopkinson, C. S., Marti, E., Packman, A. I., Newbold, J. D., & Sabater, F. (2008). Biophysical controls on organic carbon fluxes in fluvial networks. *Nature Geoscience*, *1*(2), 95–100. https://doi.org/10.1038/ngeo101

Besemer, K. (2015). Biodiversity, community structure and function of biofilms in stream ecosystems. *Research in Microbiology*, *166*(10), 774–781. https://doi.org/10.1016/j.resmic.2015.05.006.

Besemer, K., Peter, H., Logue, J. B., Langenheder, S., Lindström, E. S., Tranvik, L. J., & Battin, T. J. (2012). Unraveling assembly of stream biofilm communities. *The ISME Journal*, *6*(8), 1459–1468. https://doi.org/10.1038/ismej.2011.205.

Birk, S., Bonne, W., Borja, A., Brucet, S., Courrat, A., Poikane, S., Solimini, A., van de Bund, W., Zampoukas, N., & 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.

Borja, A. (2018). Testing the efficiency of a bacterial community-based index (microgAMBI) to assess distinct impact sources in six locations around the world. *Ecological Indicators*, *85*, 594–602. https://doi.org/10.1016/j.ecolind.2017.11.018.

Bouvier, T., & del Giorgio, P. A. (2003). Factors influencing the detection of bacterial cells using fluorescence in situ hybridization (FISH): A quantitative review of published reports. *FEMS Microbiology Ecology*, *44*(1), 3–15. https://doi.org/10.1016/S0168-6496(02)00461-0.

Bragina, A., Maier, S., Berg, C., Müller, H., Chobot, V., Hadacek, F., & Berg, G. (2012). Similar diversity of alphaproteobacteria and nitrogenase gene amplicons on two related sphagnum mosses. *Frontiers in Microbiology*, *2*, 8–9. https://doi.org/10.3389/fmicb.2011.00275.

Bunn, S. E., & Arthington, A. H. (2002). Basic principles and ecological consequences of altered flow regimes for aquatic biodiversity. *Environmental Management*, *30*(4), 492–507. https://doi.org/10.1007/s00267-002-2737-0.

Callahan, B. (2018). Silva taxonomic training data formatted for DADA2 (Silva version 132) [Data set]. *Zenodo*. https://doi.org/10.5281/zenodo.1172783.

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

Calner, M. (2013). The Lower Palaeozoic of southern Sweden and the Oslo Region, Norway field guide for the 3rd Annual Meeting of the IGCP project 591. Sveriges geologiska undersökning.

Caruso, G., La Ferla, R., Azzaro, M., Zoppini, A., Marino, G., Petochi, T., Corinaldesi, C., Leonardi, M., Zaccone, R., Fonda Umani, S., Caroppo, C., Monticelli, L., Azzaro, F., Decembrini, F., Maimone, G., Cavallo, R. A., Stabili, L., Hristova Todorova, N., K. Karamfilov, V., … Danovaro, R. (2016). Microbial assemblages for environmental quality assessment: Knowledge, gaps and usefulness in the European Marine Strategy Framework Directive. *Critical Reviews in Microbiology*, *42*(6), 883–904. https://doi.org/10.3109/1040841X.2015.1087380.

Cordier, T., Lanzén, A., Apothéloz-Perret-Gentil, L., Stoeck, T., & Pawlowski, J. (2019). Embracing environmental genomics and machine learning for routine biomonitoring. *Trends in Microbiology*, *27*(5), 387–397. https://doi.org/10.1016/j.tim.2018.10.012.

Corte, D. D., Sintes, E., Yokokawa, T., & Herndl, G. J. (2013). Comparison between MICRO–CARD–FISH and 16S rRNA gene clone libraries to assess the active versus total bacterial community in the coastal Arctic. *Environmental Microbiology Reports*, *5*(2), 272–281. https://doi.org/10.1111/1758-2229.12013.

Das Gupta, A. (2008). Implication of environmental flows in river basin management. *Physics and Chemistry of the Earth, Parts A/B/C*, *33*(5), 298–303. https://doi.org/10.1016/j.pce.2008.02.004.

Davies, P. M., Naiman, R. J., Warfe, D. M., Pettit, N. E., Arthington, A. H., & Bunn, S. E. (2014). Flow–ecology relationships: Closing the loop on effective environmental flows. *Marine and Freshwater Research*, *65*(2), 133. https://doi.org/10.1071/MF13110.

Dedysh, S. N. (2009). Exploring methanotroph diversity in acidic northern wetlands: Molecular and cultivation-based studies. *Microbiology*, *78*(6), 655–669. https://doi.org/10.1134/S0026261709060010.

Demars, B. O. L., Friberg, N., & Thornton, B. (2020). Pulse of dissolved organic matter alters reciprocal carbon subsidies between autotrophs and bacteria in stream food webs. *Ecological Monographs*, *90*(1), e01399. https://doi.org/10.1002/ecm.1399

Edgar, R. C. (2017). Updating the 97% identity threshold for 16S ribosomal RNA OTUs [Preprint]. https://doi.org/10.1101/192211.

Fazi, S., Amalfitano, S., Pernthaler, J., & Puddu, A. (2005). Bacterial communities associated with benthic organic matter in headwater stream microhabitats. *Environmental Microbiology*, *7*(10), 1633–1640. https://doi.org/10.1111/j.1462-2920.2005.00857.x.

Fazi, S., Amalfitano, S., & Pizzetti, I. (2007). Efficiency of fluorescence in situ hybridization for bacterial cell identification in temporary river sediments with contrasting water content. *Systematic and Applied Microbiology*, *30*, 463–470.

Fazi, S., Baldassarre, L., Cassin, D., Quero, G. M., Pizzetti, I., Cibic, T., Luna, G. M., Zonta, R., & Del Negro, P. (2020). Prokaryotic community composition and distribution in coastal sediments following a Po river flood event (northern Adriatic Sea, Italy). *Estuarine, Coastal and Shelf Science*, *233*, 106547. https://doi.org/10.1016/j.ecss.2019.106547.

Fazi, S., Vázquez, E., Casamayor, E. O., & Butturini, A. (2013). Stream hydrological fragmentation drives bacterioplankton community composition. Smidt H (ed.). *PLoS One*, *8*, e64109.

Freixa, A., Ejarque, E., Crognale, S., Amalfitano, S., Fazi, S., Butturini, A., & Romaní, A. M. (2016). Sediment microbial communities rely on different dissolved organic matter sources along a Mediterranean river continuum: DOM utilization by river sediments. *Limnology and Oceanography*, *61*(4), 1389–1405. https://doi.org/10.1002/lno.10308.

Friberg, N. (2014). Impacts and indicators of change in lotic ecosystems: Impacts and indicators of change in lotic ecosystems. *Wiley Interdisciplinary Reviews: Water*, *1*(6), 513–531. https://doi.org/10.1002/wat2.1040.

Gessner, M. O., & Chauvet, E. (2002). A case for using litter breakdown to assess functional stream integrity. *Ecological Applications*, *12*(2), 498–510.

Goffredi, S. K., Kantor, A. H., & Woodside, W. T. (2011). Aquatic microbial habitats within a neotropical rainforest: Bromeliads and pH-associated trends in bacterial diversity and composition. *Microbial Ecology*, *61*(3), 529–542. https://doi.org/10.1007/s00248-010-9781-8.

Gołębiewski, M., Całkiewicz, J., Creer, S., & Piwosz, K. (2017). Tideless estuaries in brackish seas as possible freshwater-marine transition zones for bacteria: The case study of the Vistula river estuary: Bacterial freshwater-marine transition in the Baltic Sea. *Environmental Microbiology Reports*, *9*(2), 129–143. https://doi.org/10.1111/1758-2229.12509.

Greuter, D., Loy, A., Horn, M., & Rattei, T. (2016). probeBase—an online resource for rRNA-targeted oligonucleotide probes and primers: New features 2016. *Nucleic Acids Research*, *44*(D1), D586–D589. https://doi.org/10.1093/nar/gkv1232.

Hallmann, C. A., Sorg, M., Jongejans, E., Siepel, H., Hofland, N., Schwan, H., Stenmans, W., Müller, A., Sumser, H., Hörren, T., Goulson, D., & de Kroon, H. (2017). More than 75 percent decline over 27 years in total flying insect biomass in protected areas. *PLoS One*, *12*(10), e0185809. https://doi.org/10.1371/journal.pone.0185809.

Hallmann, C. A., Zeegers, T., van Klink, R., Vermeulen, R., van Wielink, P., Spijkers, H., van Deijk, J., van Steenis, W., & Jongejans, E. (2020). Declining abundance of beetles, moths and caddisflies in the Netherlands. *Insect Conservation and Diversity*, *13*(2), 127–139. https://doi.org/10.1111/icad.12377.

Heiskanen, A.-S., Berg, T., Uusitalo, L., Teixeira, H., Bruhn, A., Krause-Jensen, D., Lynam, C. P., Rossberg, A. G., Korpinen, S., Uyarra, M. C., & Borja, A. (2016). Biodiversity in marine ecosystems—European Developments Toward Robust Assessments. *Frontiers in Marine Science*, *3*, 13–17. https://doi.org/10.3389/fmars.2016.00184.

Illumina, I. (2015). 16S Metagenomic Sequencing Library Preparation, Part 15044223 Rev.B.

Jackson, M. C., Weyl, O. L. F., Altermatt, F., Durance, I., Friberg, N., Dumbrell, A. J., Piggott, J. J., Tiegs, S. D., Tockner, K., Krug, C. B., Leadley, P. W., & Woodward, G. (2016). Recommendations for the next generation of global freshwater biological monitoring tools. In A. J. Dumbrell R. L. Kordas & G. Woodward (Eds.), *Advances in Ecological Research* (Vol. *55*, pp. 615–636). Elsevier. https://doi.org/10.1016/bs.aecr.2016.08.008.

Johannessen, G. S., Wennberg, A. C., Nesheim, I., & Tryland, I. (2015). Diverse land use and the Impact on (Irrigation) Water Quality and Need for Measures—A Case Study of a Norwegian River. *International Journal of Environmental Research and Public Health*, *12*(6), 6979–7001. https://doi.org/10.3390/ijerph120606979.

Johnsen, A. R., Hausner, M., Schnell, A., & Wuertz, S. (2000). Evaluation of fluorescently labeled lectins for noninvasive localization of extracellular polymeric substances in sphingomonas biofilms. *Applied and Environmental Microbiology*, *66*(8), 3487–3491. https://doi.org/10.1128/AEM.66.8.3487-3491.2000.

Johnson, S., Spakowicz, D. J., Hong, B.-Y., Petersen, L. M., Demkowicz, P., Chen, L., Leopold, S. R., Hanson, B. M., Agresta, H. O., Gerstein, M., Sodergren, E., & Weinstock, G. M. (2019). Evaluation of 16S rRNA gene sequencing for species and strain-level microbiome analysis. *Nature Communications*, *10*(1), 5029. https://doi.org/10.1038/s41467-019-13036-1.

Lee, J.-W., Nam, J.-H., Kim, Y.-H., Lee, K.-H., & Lee, D.-H. (2008). Bacterial communities in the initial stage of marine biofilm formation on artificial surfaces. *The Journal of Microbiology*, *46*(2), 174–182. https://doi.org/10.1007/s12275-008-0032-3.

Ligi, T., Oopkaup, K., Truu, M., Preem, J.-K., Nõlvak, H., Mitsch, W. J., Mander, Ü., & Truu, J. (2014). Characterization of bacterial communities in soil and sediment of a created riverine wetland complex using high-throughput 16S rRNA amplicon sequencing. *Ecological Engineering*, *72*, 56–66. https://doi.org/10.1016/j.ecoleng.2013.09.007.

Liu, H., Økland, T., Halvorsen, R., Gao, J., Liu, Q., Eilertsen, O., & Bratli, H. (2008). Gradients analyses of forests ground vegetation and its relationships to environmental variables in five subtropical forest areas, S and SW China. *Sommerfeltia*, *32*(1), 25–27. https://doi.org/10.2478/v10208-011-00126.

Liu, R., Yu, Z., Zhang, H., Yang, M., Shi, B., & Liu, X. (2012). Diversity of bacteria and mycobacteria in biofilms of two urban drinking water distribution systems. *Canadian Journal of Microbiology*, *58*(3), 261–270. https://doi.org/10.1139/w11-129.

Lupini, G., Proia, L., Di Maio, M., Amalfitano, S., & Fazi, S. (2011). CARD–FISH and confocal laser scanner microscopy to assess successional changes of the bacterial community in freshwater biofilms. *Journal of Microbiological Methods*, *86*(2), 248–251. https://doi.org/10.1016/j.mimet.2011.05.011.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *Embnet Journal*, *17*, 10.

Martínez-Santos, M., Lanzén, A., Unda, J., Martín, I., Garbisu, C., & Ruiz-Romera, E. (2018). Treated and untreated wastewater effluents alter river sediment bacterial communities involved in nitrogen and sulphur cycling. *Science of the Total Environment*, *633*, 1051–1061. https://doi.org/10.1016/j.scitotenv.2018.03.229.

Merritt, D. M., Scott, M. L., Poff, N. L., Auble, G. T., & Lytle, D. A. (2010). Theory, methods and tools for determining environmental flows for riparian vegetation: Riparian vegetation-flow response guilds. *Freshwater Biology*, *55*(1), 206–225. https://doi.org/10.1111/j.1365-2427.2009.02206.x.

Nordeidet, B., Nordeide, T., Åstebøl, S. O., & Hvitved-Jacobsen, T. (2004). Prioritising and planning of urban stormwater treatment in the Alna watercourse in Oslo. *Science of the Total Environment*, *334–335*, 231–238. https://doi.org/10.1016/j.scitotenv.2004.04.040.

Nzewi, E., Reddy, G., Luster-Teasley, S., Kabadi, V., Chang, S.-Y., Schimmel, K., & Uzochukwu, G. (Eds.). (2009). Proceedings of the 2007 National Conference on Environmental Science and Technology. Springer, New York. 47–51. https://doi.org/10.1007/978-0-387-88483-7.

Oksanen, J., Blanchet, F. R., Kindt, R., Legendre, P., Minchin, P. R., O'Hara, R. B., Simpson, G., Solymos, P., Stevens, M. H. H., & Wagner, H. H. (2013). Package 'vegan' Version 2.5-6. Retrieved from http://cran.rproject.org. The R foundation for statistical computing.

Peder Flaten, T. (1991). A nation-wide survey of the chemical composition of drinking water in Norway. *Science of the Total Environment*, *102*, 35–73. https://doi.org/10.1016/0048-9697(91)90307-Z.

Pernthaler, A., Pernthaler, J., & Amann, R. (2002). Fluorescence In Situ Hybridization and Catalyzed Reporter Deposition for the Identification of Marine Bacteria. *Applied and Environmental Microbiology*, *68*(6), 3094–3101. https://doi.org/10.1128/AEM.68.6.3094-3101.2002.

Poff, N. L. (2018). Beyond the natural flow regime? Broadening the hydro-ecological foundation to meet environmental flows challenges in a non-stationary world. *Freshwater Biology*, *63*(8), 1011–1021. https://doi.org/10.1111/fwb.13038.

Poff, N. L. R., Allan, J. D., Bain, M. B., Karr, J. R., Prestegaard, K. L., Richter, B. D., Sparks, R. E., & Stromberg, J. C. (1997). The Natural Flow Regime. *BioScience*, *47*(11), 769–784. https://doi.org/10.2307/1313099.

R Development Core Team (2008). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. http://www.R-project.org.

Reimann, C., Finne, T. E., Nordgulen, Ø., Sæther, O. M., Arnoldussen, A., & Banks, D. (2009). The influence of geology and land-use on inorganic stream water quality in the Oslo region, Norway. *Applied Geochemistry*, *24*(10), 1862–1874. https://doi.org/10.1016/j.apgeochem.2009.06.007.

Robbins, C. J., King, R. S., Yeager, A. D., Walker, C. M., Back, J. A., Doyle, R. D., & Whigham, D. F. (2017). Low-level addition of dissolved organic carbon increases basal ecosystem function in a boreal headwater stream. *Ecosphere*, *8*(4), e01739. https://doi.org/10.1002/ecs2.1739.

Salis, R. K., Bruder, A., Piggott, J. J., Summerfield, T. C., & Matthaei, C. D. (2017). High-throughput amplicon sequencing and stream benthic bacteria: Identifying the best taxonomic level for multiple-stressor research. *Scientific Reports*, *7*(1), 44657. https://doi.org/10.1038/srep44657.

Sánchez-Bayo, F., & Wyckhuys, K. A. G. (2019). Worldwide decline of the entomofauna: A review of its drivers. *Biological Conservation*, *232*, 8–27. https://doi.org/10.1016/j.biocon.2019.01.020.

Sekar, R., Fuchs, B. M., Amann, R., & Pernthaler, J. (2004). Flow sorting of marine bacterioplanktonafter fluorescence in situ hybridization. *Applied and Environmental Microbiology*, *70*(10), 6210–6219. https://doi.org/10.1128/AEM.70.10.6210-6219.2004.

Simonin, M., Voss, K. A., Hassett, B. A., Rocca, J. D., Wang, S.-Y., Bier, R. L., Violin, C. R., Wright, J. P., & Bernhardt, E. S. (2019). In search of microbial indicator taxa: Shifts in stream bacterial communities along an urbanization gradient. *Environmental Microbiology*, *21*(10), 3653–3668. https://doi.org/10.1111/1462-2920.14694.

Slagstad, T., Roberts, N. M. W., Coint, N., Høy, I., Sauer, S., Kirkland, C. L., Marker, M., Røhr, T. S., Henderson, I. H. C., Stormoen, M. A., Skår, Ø., Sørensen, B. E., & Bybee, G. (2018). Magma-driven, high-grade metamorphism in the Sveconorwegian Province, southwest Norway, during the terminal stages of Fennoscandian Shield evolution. *Geosphere*, *14*(2), 861–882. https://doi.org/10.1130/GES01565.1.

Steinnes, E., Flaten, T. P., Varskog, P., Låg, J., & Bølviken, B. (1993) Acidification status of Norwegian forest soils as evident from large

scale studies of humus samples. *Scandinavian Journal of Forest Research*, *8*(1-4), 291-304. https://doi.org/10.1080/02827589309382778.

Szabó, K. É., Ács, É., Kiss, K. T., Eiler, A., Makk, J., Plenković-Moraj, A., Tóth, B., & Bertilsson, S. (2007). Periphyton-based water quality analysis of a large river (RiverDanube, Hungary): Exploring the potential of molecular fingerprinting forbiomonitoring. *River Systems*, *17*(3–4), 365–382. https://doi.org/10.1127/lr/17/2007/365.

Tassi, F., Fazi, S., Rossetti, S., Pratesi, P., Ceccotti, M., Cabassi, J., Capecchiacci, F., Venturi, S., & Vaselli, O. (2018). The biogeochemical vertical structure renders a meromictic volcanic lake a trap for geogenic CO2 (Lake Averno. Italy). *PLoS One*, *13*(3), e0193914.

Thompson, M. S. A., Bankier, C., Bell, T., Dumbrell, A. J., Gray, C., Ledger, M. E., Lehmann, K., McKew, B. A., Sayer, C. D., Shelley, F., Trimmer, M., Warren, S. L., & Woodward, G. (2016). Gene-to-ecosystem impacts of a catastrophic pesticide spill: Testing a multilevel bioassessment approach in a river ecosystem. *Freshwater Biology*, *61*(12), 2037–2050. https://doi.org/10.1111/fwb.12676.

Tiegs, S. D., Costello, D. M., Isken, M. W., Woodward, G., McIntyre, P. B., Gessner, M. O., Chauvet, E., Griffiths, N. A., Flecker, A. S., Acuña, V., Albariño, R., Allen, D. C., Alonso, C., Andino, P., Arango, C., Aroviita, J., Barbosa, M. V. M., Barmuta, L. A., Baxter, C. V., & Zwart, J. A. (2019). Global patterns and drivers of ecosystem functioning in rivers and riparian zones. Science. *Advances*, *5*(1), eaav0486. https://doi.org/10.1126/sciadv.aav0486.

Vörösmarty, C. J. et al (2005) *Millennium ecosystem assessment*, Vol. 1, Ch. 7, pp. 165–207.Island Press.

Wang, Q., Garrity, G. M., Tiedje, J. M., & Cole, J. R. (2007). Naive bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Applied and Environmental Microbiology*, *73*(16), 5261–5267. https://doi.org/10.1128/AEM.00062-07.

Webster, N. S., Smith, L. D., Heyward, A. J., Watts, J. E. M., Webb, R. I., Blackall, L. L. & Negri, A. P. (2004). Metamorphosis of a scleractinian coral in response to microbial biofilms. *Applied and Environment Microbiology*, *70*, 1213–1221.

Widder, S., Besemer, K., Singer, G. A., Ceola, S., Bertuzzo, E., Quince, C., Sloan, W. T., Rinaldo, A., & Battin, T. J. (2014). Fluvial network organization imprints on microbial co-occurrence networks. *Proceedings of the National Academy of Sciences USA*, *111*(35), 12799–12804. https://doi.org/10.1073/pnas.1411723111.

Wilhelm, L., Singer, G. A., Fasching, C., Battin, T. J., & Besemer, K. (2013). Microbial biodiversity in glacier-fed streams. *The ISME Journal*, *7*(8), 1651–1660. https://doi.org/10.1038/ismej.2013.44.

Williams, H. G., Day, M. J., Fry, J. C., & Stewart, G. J. (1996). Natural transformation in river epilithon. *Applied and Environmental Microbiology*, *62*(8), 2994–2998. https://doi.org/10.1128/AEM.62.8.2994-2998.1996.

Williamson, M. H. (1978). The ordination of incidence data. *The Journal of Ecology*, *66*(3), 911. https://doi.org/10.2307/2259303.

Woodward, G., Gessner, M. O., Giller, P. S., Gulis, V., Hladyz, S., Lecerf, A., Malmqvist, B., McKie, B. G., Tiegs, S. D., Cariss, H., Dobson, M., Elosegi, A., Ferreira, V., Graca, M. A. S., Fleituch, T., Lacoursiere, J. O., Nistorescu, M., Pozo, J., Risnoveanu, G., … Chauvet, E. (2012). Continental-scale effects of nutrient pollution on stream ecosystem functioning. *Science*, *336*(6087), 1438–1440. https://doi.org/10.1126/science.1219534.

Zhang, M., Liu, W., Nie, X., Li, C., Gu, J., & Zhang, C. (2012). Molecular analysis of bacterial communities in biofilms of a drinking water clearwell. *Microbes and Environments*, *27*(4), 443–448. https://doi.org/10.1264/jsme2.ME12035.

Zoppini, A., Amalfitano, S., Fazi, S., & Puddu, A. (2010). Dynamics of a benthic microbial community in a riverine environment subject to hydrological fluctuations (Mulargia River, Italy). *Hydrobiologia*, *657*(1), 37–51. https://doi.org/10.1007/s10750-010-0199-6.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Pin L, Eiler A, Fazi S, Friberg N. Two different approaches of microbial community structure characterization in riverine epilithic biofilms under multiple stressors conditions: Developing molecular indicators. *Mol Ecol Resour*. 2021;21:1200–1215. https://doi.org/10.1111/1755-0998.13341