How hydropower induced changes in temperature, quality and quantity of water influences structure and functioning of riverine biofilms

Acknowledgments

This PhD thesis was part of the Euro FLOW project "EUROpean training and research network for environmental FLOW management in river basins", a European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No 765553. The project was a collaboration between 10 participants: Universidad De Cantabria (Spain), EAWAG (EIDGENOESSISCHE ANSTALT FUER WASSERVERSORGUNG ABWASSERREINIGUNG UND GEWAESSERSCHUTZ) (Switzerland), Norsk Institut For Vannforskning (Norway), The University Of Birmingham (United Kingdom), Forschungsverbund Berlin EV (Germany), Stichting IHE Delft Institute For Water Education (Netherlands), Helmholtz-Zentrum Fur Umweltforschung Gmbh – UFZ (Germany), Universita Degli Studi Di Trento (Italy), United Kingdom Research And Innovation (United Kingdom), Uk Centre For Ecology & Hydrology (United Kingdom). Among the partners of the consortium there were 15 institutions: Universitetet I Oslo (Norway), Kelda Group PLC (United Kingdom), Istituto Superiore Per La Protezione E La Ricerca Ambientale (Italy), Environment Agency (United Kingdom), Stiftung Schweizerischer National Park (Switzerland), Engadiner Kraftwerke AG (Switzerland), CBEC eco-engineering UK Ltd (United Kingdom), Landesbetrieb für Hochwasserschutz und Wasserwirtschaft (Germany), Technische Universiteit Delft (Netherlands), Slovenian Water Agency (Slovenia), Ove Arup & Partners International Limited (United Kingdom), Asociacion BC3 Basque Centre For Climate Change - Klima Aldaketa Ikergai (Spain), Stantec UK Limited (United Kingdom).

I would like to thank my group of supervisors - Nikolai Friberg (NIVA; University of Copenhagen), Leif Asbjørn Vøllestad (UiO), Therese Fosholt Moe and Alexander Eiler (UiO). It has been an amazing experience to work with you all. You were always there when I needed help and you always came up with brilliant suggestions and comments, especially by the end of the PhD, when COVID-19 made everything more challenging.

I would like to thank Nikolai for his insightful discussions and for always being positive, even in the toughest moments. I would like to thank Alex for his help with everything concerning coding and microbial related issues, if it wasn't for you I would have never made it to the end.

I would like to thank Asbjørn and Therese for the strong commitment during the writing of the papers.

I would like to thank all the colleagues at NIVA, which made the endless days of working a bit lighter with all the chats and laughs.

A special thanks goes to Liora, for always being on my side through this rollercoaster of emotions which is the PhD and, for never giving up on me even when I was gloomy and sad.

Thanks for being part of my life.

Contents

Summary 7
List of papers 10
Declaration of contribution 11
Authors contributions 11
Introduction 13
The ecology of river ecosystems 13
Impact of hydropower development and flow alterations in river systems 15
Prokaryotic biofilm communities in river systems 17
Norway as case study 21
Aims 23
Methods and results 24
Paper I 24
Paper II 30
Paper III 35
Paper IV 41
Discussion 46
Molecular methodologies for the analysis of prokaryotic community structure in Norway (Paper I) 46
Response of prokaryotic community structure and functioning to modified flow regimes (Paper II-III) 49
Hydropower impact in Norwegian river systems 49
Response to different flow regime scenarios 52
Developing prokaryotic bioindicators for classification of the ecological status of river ecosystems (Paper I-IV) 55
Perspectives 57
References 58
Websites 71

Summary

River ecosystems are hot-spots of biodiversity, hosting many unique species and providing key ecosystem services. Nowadays these ecosystems are threatened by multiple stressors and human impact is the paramount reason. Flow alterations due to hydropower development and damming are among the most important sources of impact in river ecosystems, having detrimental effects on their fragile ecology. This thesis will look at a biological component so far neglected by the research addressing the effects of flow alteration on freshwater organisms, namely the prokaryotes.

Prokaryotes are key players in the biogeochemical processes of river ecosystems. Most prominently, they have a variety of adaptive mechanisms which makes them capable of colonizing almost every ecological niche. Within river systems, prokaryotes are mainly found in biological matrices colonized by multiple species, also called biofilms. Biofilms were the focus of this thesis (**paper I-II-III**), both attached to inorganic and organic substrates. In **paper I** a combination of quantitative (Catalyzed Reported Deposition -Fluorescence in situ Hybridization) and semi-quantitative (16S rRNA metabarcoding) molecular techniques was used to screen the biofilm communities living in Norwegian river systems affected by multiple stressors. Strong environmental gradients were driving the prokaryotic community composition, revealing the importance of features such as pH and nutrient loads for the community dynamic. In this study a new methodology to detect prokaryotic indicator taxa by using metabarcoding data is presented. Taxa showing highest variance and prevalence across the sampling sites were selected as candidate bioindicators. Specific relationships between the candidate bioindicators (at different taxonomic levels) and the main environmental drivers were detected. The impact of hydropower and dams on the prokaryotic community structure and functioning was assessed in **paper II** by using 16S rRNA gene metabarcoding and litter bags with different mesh sized to analyze the

organic matter breakdown in impacted and unimpacted reaches of ten Norwegian river systems. The biofilm community structure showed no clear pattern when looking at upstream downstream gradients, while the organic matter breakdown seemed to be significantly influenced by the presence of the dams, with the downstream reaches showing higher decomposition.

To further analyze the effects of flow regime alterations on the prokaryotic communities and the biogeochemical processes they carry out, in **paper III,** we designed a mesocosm experiment where sixteen stainless steel flumes were used to implement four different flow regimes scenarios typical of managed river systems: homogenized flow regimes; flow regimes affected by agricultural management; drought treatments with water level reduced by 60%; and natural flow regime as control. Major effects of flow alterations on the prokaryotic community structure and functioning were detected. Drought treatments showed the strongest impact on both the prokaryotic assemblages and the decomposition of organic matter, which was significantly higher compared to the other treatments. Significant differences in the beta diversity were found among the prokaryotic communities from the four treatments, however we did not always find differences in the organic matter breakdown. In **paper IV** we employed machine learning algorithms to detect prokaryotic bioindicators for water quality classification in the Danube river by using an existing dataset of 16S metabarcoding for planktonic microbes. We found that different sequence variants combinations for the microbial community in upstream sites were yielding accurate prediction for water quality classification at a given downstream site. A number of sequence variants were found to be functionally redundant and, as such, mutually exclusive in terms of predictive information. These mutually exclusive sequence variants were at times phylogenetically related, and thus we can conclude that specific taxonomic groups have a generalizable predictive power and would need to be tested in other river basins to evaluate their performance in predicting water quality.

The results from this thesis showed significant changes in both structure and functioning of prokaryotic communities in relation to flow alterations due to river management, damming, water

abstraction and hydropower developments. However, it was also evident that effects on prokaryotic communities were very context dependent with catchment specific differences sometime overruling effects of flow changes in terms of community composition. The clear response found in the experimental study of flow alterations supports the contention that natural variability is high when undertaking field studies and is masking singular effect of flow. Overall, the findings in this thesis suggest a potential and widely overlooked impact of flow alterations on riverine microbial communities. Shifts in the prokaryotic assemblages might have repercussions on the energy transfer within the freshwater food webs, impairing crucial ecosystem processes such as carbon and nitrogen cycling, which is extremely relevant considering future climate change scenarios. By showing that identification of prokaryotic indicators for water quality classification is possible using various statistical tools, this thesis suggests that prokaryotic indicators can be implemented in current monitoring networks. The usefulness of these indicators is further exemplified providing new insights in the ecological status of river ecosystems impacted by human induced stressors such as hydro-morphological alterations due to hydropower and damming.

List of papers

Paper 1.

Pin, L., Eiler, A., Fazi, S., & Friberg, N. (2021). Two different approaches of microbial community structure characterization in riverine epilithic biofilms under multiple stressors conditions: Developing molecular indicators. Molecular Ecology Resources, 1755-0998.13341. https://doi.org/10.1111/1755-0998.13341

Paper II.

Pin, L., Eiler, A., Vøllestad, L. A.., Moe, T. & Friberg, N. Impact of hydropower at the base of river food webs: prokaryotic community composition of epilithic biofilms and organic matter breakdown. submitted to *Aquatic Ecology. 25-10-2021*

Paper III.

Pin, L., Arias Font, R., Ledger, M. E., Eiler, A., Vøllestad, L. A. & Friberg, N. Flow regimes determine microbial community structure and functioning: a mesocosm approach. Manuscript to be submitted to *Science of the Total Environment*

Paper IV.

Fontaine, L., **Pin, L.**, Savio, D., Farnleitner, A., Kirschner, A., Friberg, N., Eiler, A. Prokaryotic bioindicators for water quality along a continental river. Manuscript to be submitted to *ISME journal*

Declaration of contribution

Authors contributions

Lorenzo Pin (LP)

Nikolai Friberg (NF)

Alexander Eiler (AE)

Leif Asbjørn Vøllestad (AV)

Therese Fosholt Moe (TFM)

Stefano Fazi (SF)

Raquel Arias Font (RAF)

Mark E. Ledger (MEL)

Laurent Fontaine (LF)

Domenico Savio (DS)

Andreas Farnleitner (AF)

Alexander Kirschner (AK)

Paper 1

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

Paper 2

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

Paper 3

LP and NF, MEL, RAF conceived the study and the experimental design. MEL, RAF provided the mesocosm facility and collected all the water chemistry parameters and flow measurements. LP and RAF worked on the sampling and data collection. LP analyzed the data. LP, AE, AV, MEL RAF and NF discussed the results and wrote the manuscript.

Paper 4

LF, LP, AE, DS, AF conceived the study. DS, AK, AF collected the samples and processed the data. LP, DS, AE, LF worked on the data analysis. LF worked on the machine learning model development. LP worked on the multivariate statistics. LF, AE, DS, AF, AK analyzed the results and worked on the manuscript.

Introduction

The ecology of river ecosystems

Freshwater environments cover only 2.3% of Earth land surface and only 0.01% of the water globally available is actually fresh water, excluding glaciers (Dudgeon et al., 2006; Reid et al., 2019). Despite the infinitesimal area covered by freshwater ecosystems, these are among the richest environments on Earth in terms of biodiversity, accounting for around 9.5% of all animals species (Zarfl et al., 2019).

River ecosystems in particular, hold a central role in the water cycle and in the global flux of nutrients, being landscape receivers and conveying the energy sources they carry to the marine environment, sustaining therefore multiple ecosystems along their path and influencing several biological communities at the same time (Jungwirth et al., 2000; Dudgeon, 2019).

At the base of river food webs, we find primary production as the main source of energy. Green algae, diatoms and cyanobacteria are among the most important primary producers of river systems. These autotrophic microorganisms can be found in many different habitats, living both, within biofilms, attached on organic and inorganic surfaces, and in the water column as plankton (Allan et al., 2021). Another important source of energy entering in river systems is leaves, wood and animal organic matter coming from the riparian zones, also known as the brown web (Kaspari, 2004) and the areas surrounding the channel. This is often regarded as allochthonous organic matter, in contrast with the autochthonous organic matter produced directly within the channel by autotrophic microorganisms and aquatic plants (Hein et al., 2003). The finest part of the detritus entering the rivers, also known as Dissolved Organic Matter (DOM), is immediately transformed by bacteria, fungi and protists, which are key components of the “Microbial loop”, a fundamental process in aquatic ecosystems (Meyer, 1994). Thanks to these decomposers, DOM is converted into microbial biomass which is biologically available for the higher trophic levels.

Among the primary consumers in river ecosystems we find macroinvertebrates such as insect larvae, crustacean, mollusks and numerous other *taxa*. These organisms have several feeding traits, ranging from grazers, shredders, filter feeders and predators. Usually, the main source of food for the macroinvertebrates are biofilm communities, ubiquitous in the river substrates and composed by prokaryotes, algae, fungi and protists (Allan et al., 2021). Higher trophic levels in river food webs are composed by fish, amphibians, birds and mammals and, these organisms can feed on a wide array of resources, ranging from algae, invertebrates, organic detritus and to vertebrates.

The diversity of species in river ecosystems is directly dependent on the diversity of habitats available at different spatial scales. Multiple habitat types translate into more ecological niches to be colonized; thus, the biodiversity of river systems relies on the hydro-geomorphological forces shaping the channel. We could describe the river ecosystem as a continuum, as conceptualized by Vannote et al. (1980), where energy inputs change longitudinally, along with the composition of biological communities, determining different ratios between primary production and ecosystem respiration. Usually we see low ratios of primary production to ecosystem respiration (P/R) in the headwaters. These river reaches are often shaded, under natural conditions, by a dense tree canopy limiting algal growth and the main energy source is allochthonous organic matter derived from leaves and wood. In contrast, higher-order rivers are usually so wide that shading and the overall contribution of terrestrial derived carbon are diminished compared with headwaters, while autochthonous production will, in theory, increase in importance. However, many larger rivers are also deep and turbid, preventing the in-stream primary production to prevail over heterotrophic metabolism based on microbial activity. The P/R ratio is usually higher than 1 in the middle reaches of the rivers, where the canopy cover is not shading the riverbed, and the water is relatively shallow and clear enough to facilitate autotrophy by benthic organisms (both, algae and macrophytes).

River systems are defined healthy, when the majority of the water flowing in the channel keeps its natural variability and quantity along with the sediment transported and nutrient concentrations remaining within their background/undisturbed ranges. Habitat diversity and biodiversity are tightly linked and crucial to maintain the ecosystems balance as, all the biogeochemical processes happening in the river, are strictly dependent on the energy exchange among biological communities and environment. Therefore, preserving the natural flow variability is of primary importance as it is the main driver that keeps river ecosystems safe and sound.

Impact of hydropower development and flow alterations in river systems

Freshwater ecosystems, rivers in particular, have always been extremely important for human society, as they provide many ecosystem services, among which food, drinking water and energy production are essentials (Wohl, 2018). However, river ecosystems are facing an uncertain future, because of the threats posed by humans which are leading towards an exponential degradation, mainly caused by damming, water diversion and abstraction, channelization and climate change (Palmer & Ruhi, 2019).

Among the most insidious sources of impact on river ecosystems we have hydropower developments. Often regarded as the most important among the "green" energy sources nowadays, hydropower accounts for about ¾ of the total energy production from renewables (Zarfl et al., 2019). However, hydroelectric energy comes at a cost, today around 15% of the global annual runoff is stored in more than 45 000 large dams (>15 m) (Nilsson et al., 2005). In rivers impounded for hydropower purpose, quantity, timing and variability of water flow are disrupted and far from natural flow regimes (Poff et al., 1997; Aristi et al., 2014). Usually, dams enhance the flow stability in the downstream reaches, reducing the normal occurrence and magnitude of floods. Reduced flow variability is translated into lower habitat diversity within the river and in the riparian zone (Ponsatí et al., 2015). Hydrological and morphological variability of

the river are crucial to sustain biodiversity and biogeochemical processes, while systems impacted by hydropower, often have their water quality dampened and become populated only by generalist species, capable of adapting to pollution or drought (Palmer & Ruhi, 2019). Dams can also alter the sediment transport downstream, trapping large amounts of fine particles and promoting erosive processes downstream, which result in a pronounced channel incision and stabilization (Batalla et al., 2004). Another major issue caused by dams is the release of hypolimnetic water for energy production, which can revert the normal thermal regime of the river, cooling down the downstream reaches in summer and warming up during winter. Altered thermal regimes can have a huge impact on the biological cycle of certain species, affecting growth rates, mating success or metabolic rates (Prats et al., 2010). The water contained in reservoirs can often become stratified and the deeper layers can become anoxic, due to the absence of gas exchange with the surface. Anoxic conditions in hypolimnion can impact the redox status, promoting reduction of compounds such as Fe (II), NH_4^+, H_2S, which can be then released downstream where they can be oxidized, reducing the capacity of the lower river reaches to oxidize pollutants or organic matter (Ponsatí et al., 2015).

Climate change is an additional issue to be considered when assessing the potential effects of hydropower on river ecosystems. Increases in temperature and carbon dioxide concentrations are likely to have a significant impact on the biogeochemical processes in river ecosystems. The autotrophic community, for example, may shift from eukaryotic algae to a dominance of cyanobacteria which may cause severe toxic blooms (Visser et al., 2016).

The detrimental effects of altered flow regimes has been extensively studied for biological communities such as fish (Dudgeon et al., 2006; Richter et al., 2010; Wang et al., 2018; Reid et al., 2019), macroinvertebrates (Meißner et al., 2018; Sabater et al., 2018), algae (Smolar-Žvanut & Mikoš, 2014; Sabater et al., 2018) and macrophytes (Bejarano & Sordo-Ward, 2011; Nunes & Adams, 2014; Rivaes et al., 2015). On the other hand, there is a lack of knowledge regarding the

effects of damming, hydropower development and flow alteration on the prokaryotic communities, which is the focus of this PhD project.

Prokaryotic biofilm communities in river systems

Prokaryotes have a pivotal role in freshwater ecosystems, mediating several biogeochemical processes and being the main food source at the base of the riverine trophic chain (Besemer, 2015; Battin et al., 2016). They are crucial for the nutrient cycling, as they take part in degradation of organic matter and mineralization of nutrients, making it available for primary producers, and at the same time microbes themselves become food for the higher trophic levels, completing what is known as microbial loop (Besemer et al., 2005).

These organisms are mostly found in multi-cellular and multi-species consortia called biofilms (Fig. 1), the main microbial form of life in aquatic ecosystems (Besemer, 2015). Biofilms typically develop on benthic inorganic and organic surfaces (Findlay et al., 2002). The community structure of freshwater benthic biofilms is strictly dependent on the chemical composition of the substratum and show high sensitivity to changes in water quality (Romaní et al., 2013; Burgos-Caraballo et al., 2014; Battin et al., 2016; Sabater et al., 2016). Typically, freshwater biofilms (**Fig. 1**) are hot spots of biodiversity where we can find organisms belonging to the three domains of life. Light is the main driver determining the species composition and the dominance relationships among autotrophs and heterotrophs (Battin et al., 2016). Autotrophic organisms such as eukaryotic algae (among which diatoms, green algae, red algae and other) and cyanobacteria are among the main biofilm components in particular in the euphotic zone of the rivers, whereas bacteria and archaea are prevalent in the aphotic layers (Battin et al., 2016). Other important inhabitants of biofilms are fungi and protozoans and to some extent higher organisms such as insect larvae and invertebrates. These are the top consumers in the biofilm community, and their activity as grazers can be crucial for the species composition, physical structure and

nutrient cycling (Lawrence et al., 2002; Dopheide et al., 2008; Böhme et al., 2009; Risse-Buhl et al., 2012; Wey et al., 2012).

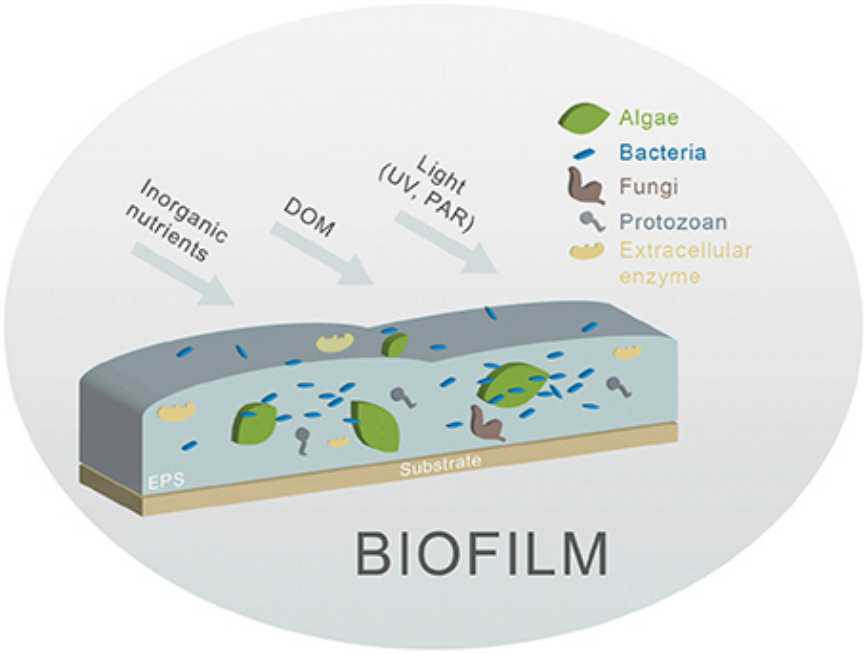


Figure 1. Schematic overview of a freshwater biofilm (Sabater et al., 2016)

The prokaryotic domain is a major component of freshwater biofilms in terms of both biomass and functional processes (Battin et al., 2016). Among the dominant prokaryotic *taxa* colonizing the benthic biofilms we find Proteobacteria and Bacteroidetes (Besemer et al., 2012; Wilhelm et al., 2013; Battin et al., 2016). Proteobacteria are the most important group of bacteria across many freshwater environments, Alphaproteobacteria and Betaproteobacteria are among the main classes dwelling in river ecosystems (Battin et al., 2016; Besemer et al., 2012). The former is an ubiquitous taxon, encompassing the members of the SAR11 clade, regarded as the most abundant taxon in marine bacterioplankton and likely to play an important role in the processing of organic matter in aquatic ecosystems (Azam & Malfatti, 2007; Eiler et al., 2009). The latter, instead, are typically found in biofilms and plankton of freshwater environments and, some groups such as the annamox, are crucial players in the freshwater nitrogen cycle (Newton et al., 2011; Jezbera et al., 2012; Salcher et al., 2016).

Among the Bacteroidetes phylum, Flavobacteria and Sphingobacteriia are instead classes which are regarded to have a central role in stream biofilms because of their ability to break down complex biopolymers such as chitin and cellulose. Their role in degrading high molecular weight compounds, which are usually the most refractory fraction of the organic matter, and their capacity of resisting to high flows, may be crucial for the correct functioning of the organic matter cycling in the river ecosystem (Kirchman, 2002; Newton et al., 2011). Other important taxa for freshwater biofilms belong to Gammaproteobacteria, Actinobacteria, Firmicutes, Deltaproteobacteria, Gemmatimonadetes, Verrucomicrobia, Planctomycetes and, recent studies revealed that also Archaea may play a role in anoxic layers (Besemer et al., 2012; Wilhelm et al., 2013).

In river ecology, the role of microbial communities is often only addressed in studies on ecosystem functioning, typically by using leaf litter bags for decomposition assessment (Gessner & Chauvet, 2002; Woodward et al., 2012; Tiegs et al., 2019). Exploring the diversity of microbial communities, and the functional aspects related to the key ecosystem processes mediated by them, will be an important asset to develop our understanding about river ecosystem functioning (Besemer, 2015).

The small dimension and high surface/volume ratio of these microorganisms makes them extremely sensitive to changes in the concentration of nutrients or pollutants (Martínez-Santos et al., 2018). This characteristic might be an advantage, which could render prokaryotes the perfect candidates as bioindicators and early warnings, useful to detect any sign of stress in the environment (Bloem & Breure, 2003).

Historically, microbes have always had less emphasis in ecological studies, mainly because of the lack of proper tools to inspect and analyze the diversity of these microscopic organisms, which often could not be cultivated in laboratories (Prosser et al., 2007). During the past decades, new molecular techniques, to study microbes in natural environment, have been developed. Among these we find direct cell-counts methods such as hybridization in situ (Fluorescence in situ

hybridization (FISH) and Catalyzed Reported Deposition -FISH (CARD-FISH), Pernthaler et al., 2002), flow-cytometry (Dann et al., 2017), quantitative Polymerase Chain Reaction (qPCR) (Thompson et al., 2016) and advanced techniques as 16S rRNA amplicon sequencing (Johnson et al., 2019), which allow us, to finally get an insight in the diversity and density of microbial species. Although several advances have been made and the molecular techniques available nowadays allow for extensive studies on microbial communities, the current environmental monitoring networks such as the Water Framework Directive (WFD), prokaryotes and microbes in general, have been neglected (Heiskanen et al., 2016; Cordier et al., 2019).

Taxonomic groups such as prokaryotes and protists have been ignored, but they might bring valuable information to the environmental impact assessment frameworks, as they are known to be very sensitive to environmental changes. New High Throughput Sequencing technologies coupled with metabarcoding have been suggested to have huge potential in bioassessment, including information from biological communities such as prokaryotes so far poorly addressed. Among prokaryotes, only cyanobacteria are currently used as bioindicators (Mateo et al., 2015), but metabarcoding could open up to the whole prokaryotic community composition, including heterotrophic bacteria and archaea. Some efforts have been already made to promote the use of prokaryotes as bioindicators, proving them to be efficient in detecting effects of multiple stressors and anthropogenic impacts in many parts of the world (Fortunato et al., 2013; Aylagas et al., 2017; Salis et al., 2017; Borja, 2018). Although there are still many limitations to these methods, relating to both, biological and technical issues (Pawlowski et al., 2018), the current development of molecular methods has already led to faster and more efficient ecological assessments, complementing the morphological taxonomic identification methods. However, molecular methodologies and bioinformatic processing of sequencing data would still need to be standardized across countries and laboratories to have comparable results, at the current stage there is still no common agreement (Cordier et al., 2020).

Norway as case study

Norway is a global leader when it comes to hydropower development, in fact more than 95% of the total electricity production of the country comes from hydroelectric energy sources (https://energifaktanorge.no/). More than 1,650 hydropower developments provide energy for the whole country, in addition, there are more than 1,000 reservoirs with a storage capacity corresponding to

70% of the total annual energy consumption in Norway (https://energifaktanorge.no/) (**Fig. 2**).

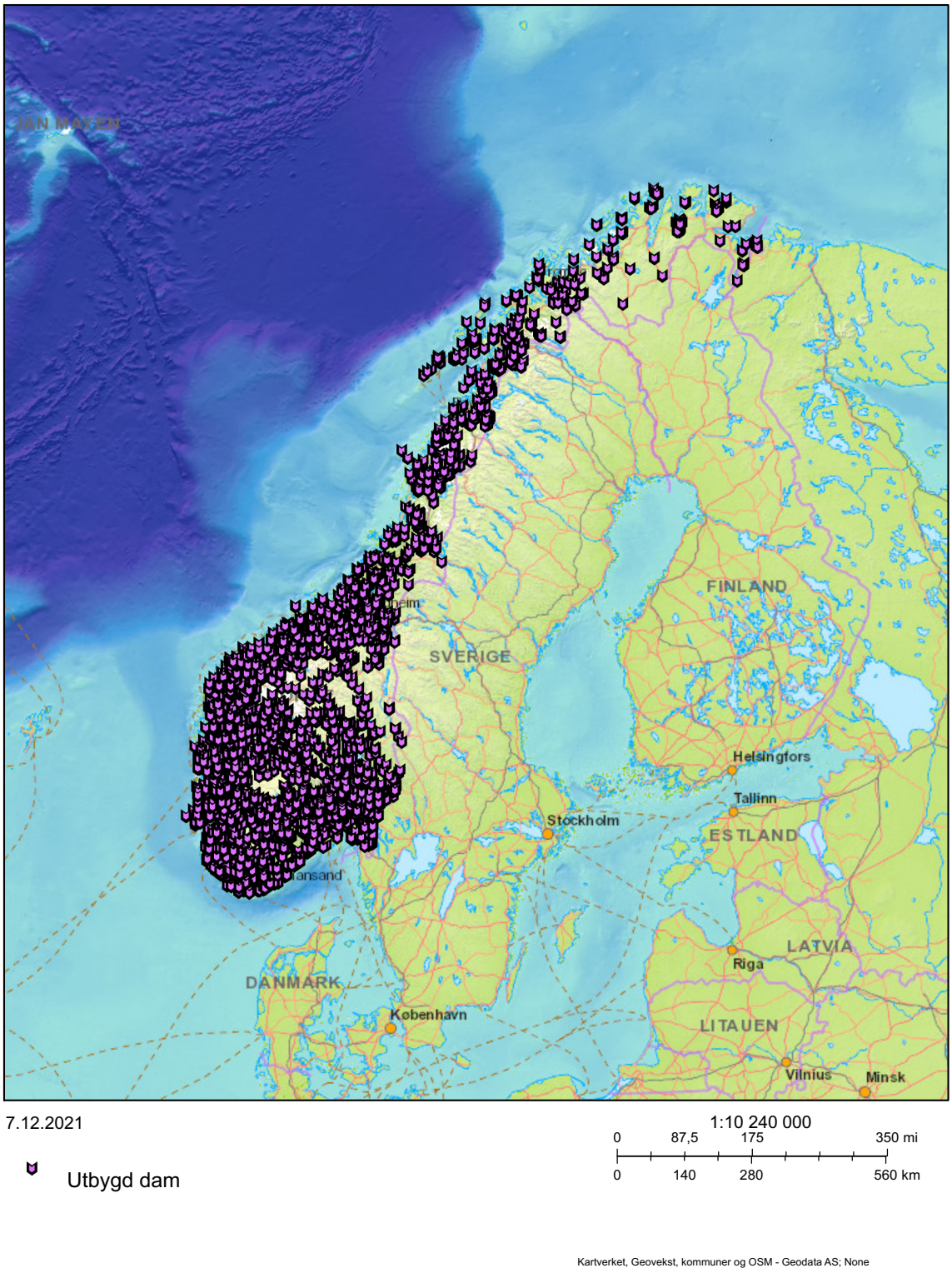


Figure 2. Map of the dams built up to day in Norway. (source https://atlas.nve.no/)

In total 2/3rd of Norwegian river systems are impacted by hydro-morphological alterations (Schneider et al., 2018).

The large number of hydropower developments have made Norway a perfect case study to address the impact of altered flow regimes and river fragmentation on river ecology. So far, many research projects have been carried out, trying to understand the consequences of modified flow regimes on the river biota, mainly focusing on fish (Thorstad et al., 2003; Bakken et al., 2012; Wright et al., 2017; Dorber et al., 2019) and aquatic macrophytes (Moe et al., 2013; Schneider et al., 2013). Overall, there is still a research gap concerning the effects of alterations to the natural flow regime at the base of the riverine food webs, especially regarding microbial communities, which are still a black box in terms of consequences on the ecology of river systems.

Aims

The aims of this study were:

1. Characterization of the microbial community structure in Norwegian river systems using different molecular tools (**paper I and II**).
2. Assess the impact of different drivers (environmental and human induced) for the structuring and functioning of biofilm prokaryotic communities (**paper I, II and III**).
3. Study the impact of flow alterations on prokaryotic community structure and organic matter breakdown in different scenarios, related to climate change and river management (**paper II and III**)
4. Explore multiple methodologies to develop new bioindicators for environmental impact assessment on river ecosystems using prokaryotic taxonomic data (**paper I and IV**).

Methods and results

Paper I

In **paper I** rivers which were either free-flowing or impacted by hydropower were selected considering areas ranging from human impacted to natural systems (**Fig. 3**).

Figure 3. Some of the hydropower facilities and dams impacting the rivers object of the first study. The development types ranged from old concrete storage reservoirs to newer run-of-the-river powerplants (Photos: Lorenzo Pin).

To characterize the microbial community of the epilithic biofilms in the selected rivers, two molecular techniques were employed, a quantitative molecular method, the Catalyzed Reported Deposition Fluorescence in situ Hybridization and a semi-quantitative method, namely 16S rRNA gene amplicon sequencing.

The former methodology is based on the use of specific oligonucleotidic probes labelled with rRNA-target horseradish peroxidase (HRP) which, reacting with a fluorescent compound, make

the prokaryotic cells glow. Thanks to the use of the epifluorescence microscope, it is possible to count the actual number of prokaryotic cells related to the specific *taxa* targeted with the probe. By comparing the number of cells identified by the probe and the total prokaryotic abundances, identified by using the 4′-6-diamidino-2-phenylindole (a general marker for all the nucleic acids, also known as DAPI), we can get an idea of what the prokaryotic community composition looks like (**Fig. 4**).

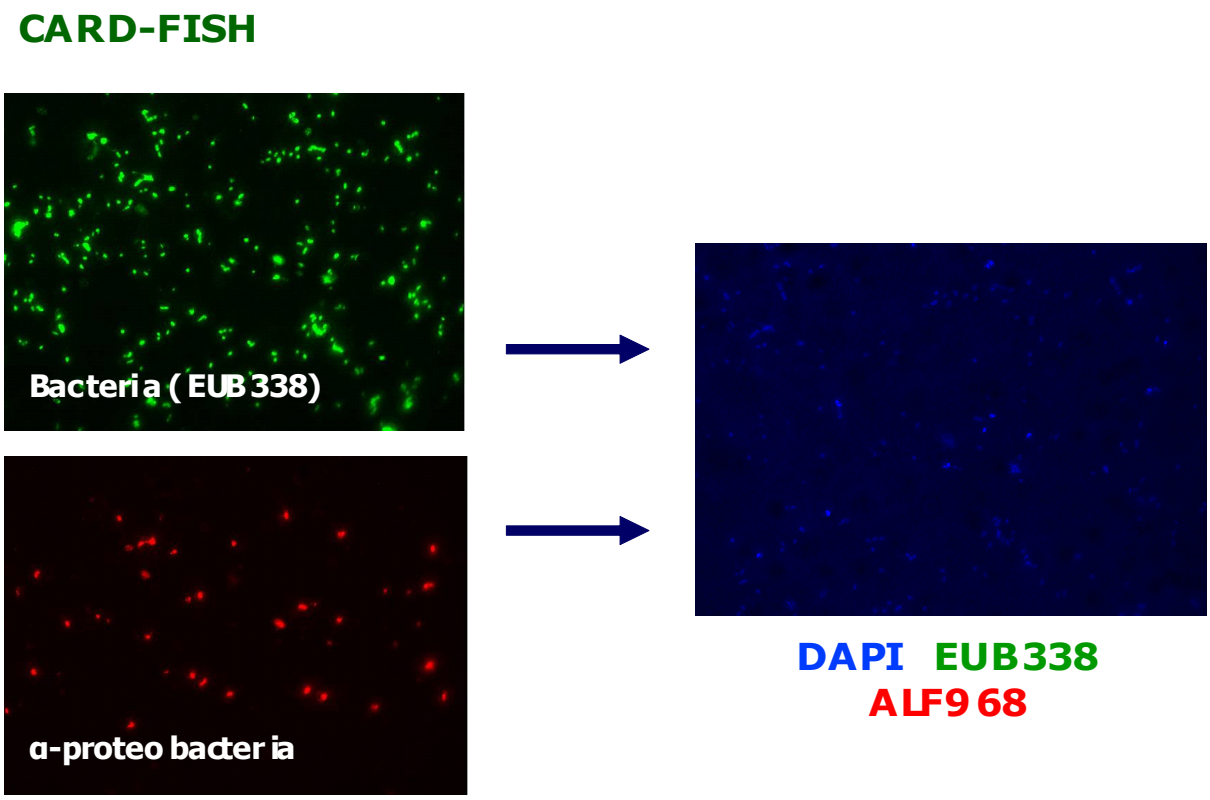


Figure 4. Epifluorecence microscope photos of prokaryotic cells hybridized with the general probe for Bacteria (Eub338) in green, with the probe specific to the Alphaproteobacteria (ALF968) in red and, in blue we can see the total abundance of the prokariotic cells marked with DAPI

I used probes targeting the domains of Archaea and Bacteria, furthermore the bacterial community was analyzed by using specific probes for the classes of the Proteobacteria phylum including Alpha-, Beta-, Gamma- and Deltaproteobacteria, and the Firmicutes phylum. Sequencing of 16S rRNA amplicons, was identified as the preferred option for the bacterial analyses, and thus used in **papers II** and **III**. Amplicon sequencing of the 16S rRNA gene is nowadays the paramount method used to assess the composition of microbial communities.

In this study, the 16S rRNA gene was used as phylogenetic marker to classify bacteria and archaea, taxonomically. This gene is highly conserved but variable at the same time and resistant to horizontal gene transfer, making it suitable for phylogenetic classification up to genus level. Typically, regions of the 16S are targeted by using specific primers and the gene is amplified through Polymerase Chain Reaction (PCR), then the amplified DNA fragments, the amplicons, can be directly sequenced through different types of technologies and the output data are then processed through bioinformatic pipelines to analyze the microbial community composition, analyze alpha and beta diversity (**Fig. 5**).

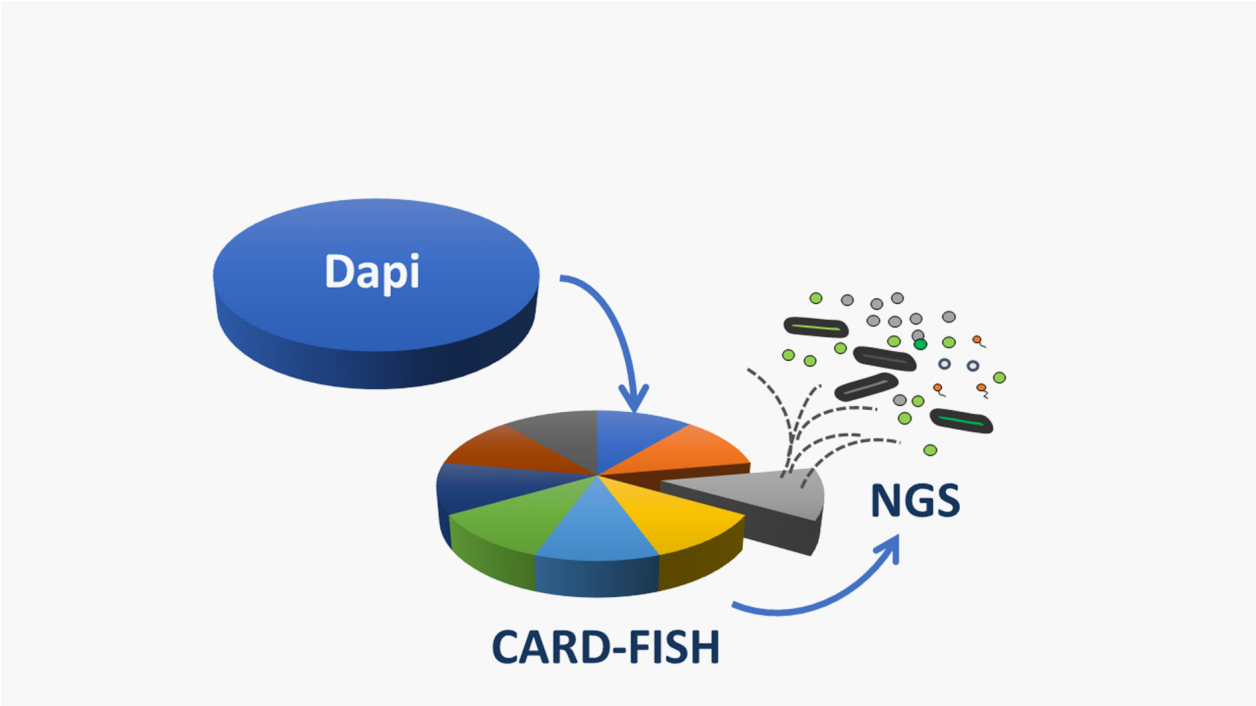


Figure 5. Conceptual schematization of the taxonomic resolution for hybridization in situ and metabarcoding approaches used in this study to analyze the prokaryotic community structure.

The results from the two techniques were compared to analyze their efficacy in describing the microbial community composition. By collecting also physico-chemical parameters of the water it was possible to analyze the patterns between environmental gradients and biological data retrieved by using the two molecular techniques. The sequencing data were also used to detect possible biological indicators among the prokaryotic Amplicon Sequence Variants (ASVs*)*

identified. The ASVs were clustered at three taxonomic levels (phylum, class and genus) and, for each of the levels, the *taxa* were selected sorting them according to their prevalence among the sampling sites and their variability.

The rationale behind this method, is that *taxa* with high prevalence (occurring in more sampling locations) and high variance (high variability of abundance) may be suitable as biological indicators for specific environmental gradients. Therefore, the *taxa* with highest prevalence and variance were used in a distance-based Redundancy Analysis in addition to the environmental variables found to be the best descriptors for the sampling sites distribution. This method allowed to identify relationships between environmental variables and specific *taxa* regarded as candidate bioindicators.

Overall, the main variance in the data could best be due to differences between the two geographical regions, south western (Agder) and south eastern (Oslo) Norway, in particular differences in the pH (respectively 5.70 ± 0.40 and 7.53 ± 0.42 on average). The main reason explaining this difference is related to the geology of the riverbed characterized by magmatic and metamorphic rocks in south western Norway vs. limestone in the south east. A similar pattern as for pH was also observed for the main nutrients such as Total Organic Carbon (TOC), Total Nitrogen (TN) and Total Phosphorus (TP), showing lower concentrations in the south west and higher in the south east.

Significant relationships in terms of R squared values were found by using a Mantel Test which compared the Bray-Curtis matrices for the prokaryotic *taxa* from both, hybridization in situ and metabarcoding to a Euclidean matrix for the environmental variables. This close correlation confirmed the strong influence of the abiotic features on the microbial community composition and abundances.

The CARD-FISH results allowed to quantify the abundance of the selected prokaryotic taxa, highlighting the relationships between dominant taxa. When analyzing the abundance of the prokaryotic taxa in the rivers impacted by human and hydropower, it was possible to identify

different ratios between the two most abundant taxa, the Alphaproteobacteria and the Gammaproteobacteria. The former class was dominant in the rivers from Agder, instead Gammaproteobacteria were co-dominating the community in the rivers from Oslo. This different behavior among these taxa may be related to the different nutrient availability of the two regions. The rivers from Agder are consistently more oligotrophic when compared to the rivers from Oslo, which suffer a considerable urbanization and higher nutrient loads (**Fig. 6**).

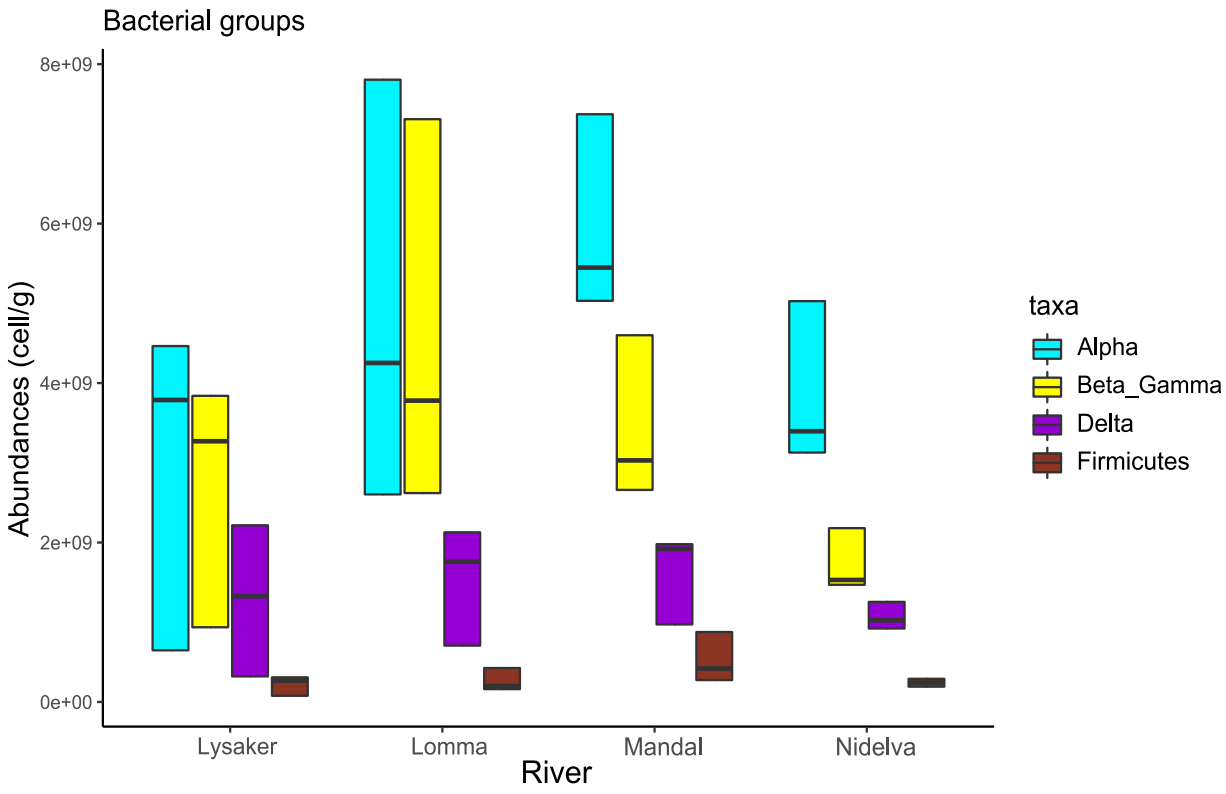


Figure 6. Cell abundances for the four taxonomic groups analyzed through CARD-FISH. Results reveal that rivers from Oslo (Lysaker and Lomma) show a codominance of Alpha- and Gammaproteobacteria, while rivers from Agder (Mandal and Nidelva) are dominated only by the Alphaproteobacteria.

From the sequencing data it was possible to reconstruct the taxonomic composition of the sampling sites down to *genus* level. Proteobacteria were found to be the most abundant phylum, confirming the results obtained by using the CARD-FISH. The results of the *envfit* analysis, which were used to test the correlation between selected environmental variables and the Bray-Curtis dissimilarity matrix of the ASVs, confirmed the significant relationship between pH and

the microbial community structure ($r = 0.9$, $p < 0.001$). The Bray-Curtis matrix for the sequencing data, similarly to the environmental variables, showed high positive correlation with the source catchment (Analysis of Similarity; $r = 0.8$, $p < 0.001$), confirming the high level of dependence of microbial community composition from the provenience of the headwaters.

Among the taxa identified by the Redundancy Analysis on the ASVs at different taxonomic levels, some were significantly related to specific environmental drivers such as *Acinetobacter*, positively related to higher values of pH, *Janthiniobacterium* with high concentrations of TP and *Sphingomonas* with low values of TN, TP and pH. These *taxa* were among the most prevalent, occurring at each sampling site and, they were also among the most variable in terms of abundance (**Fig. 7**). These bacteria might be well suited as biological indicators given their response to specific drivers, further analyses are needed to test their efficiency in predicting specific environmental parameters and thus become useful tools for environmental impact assessment.

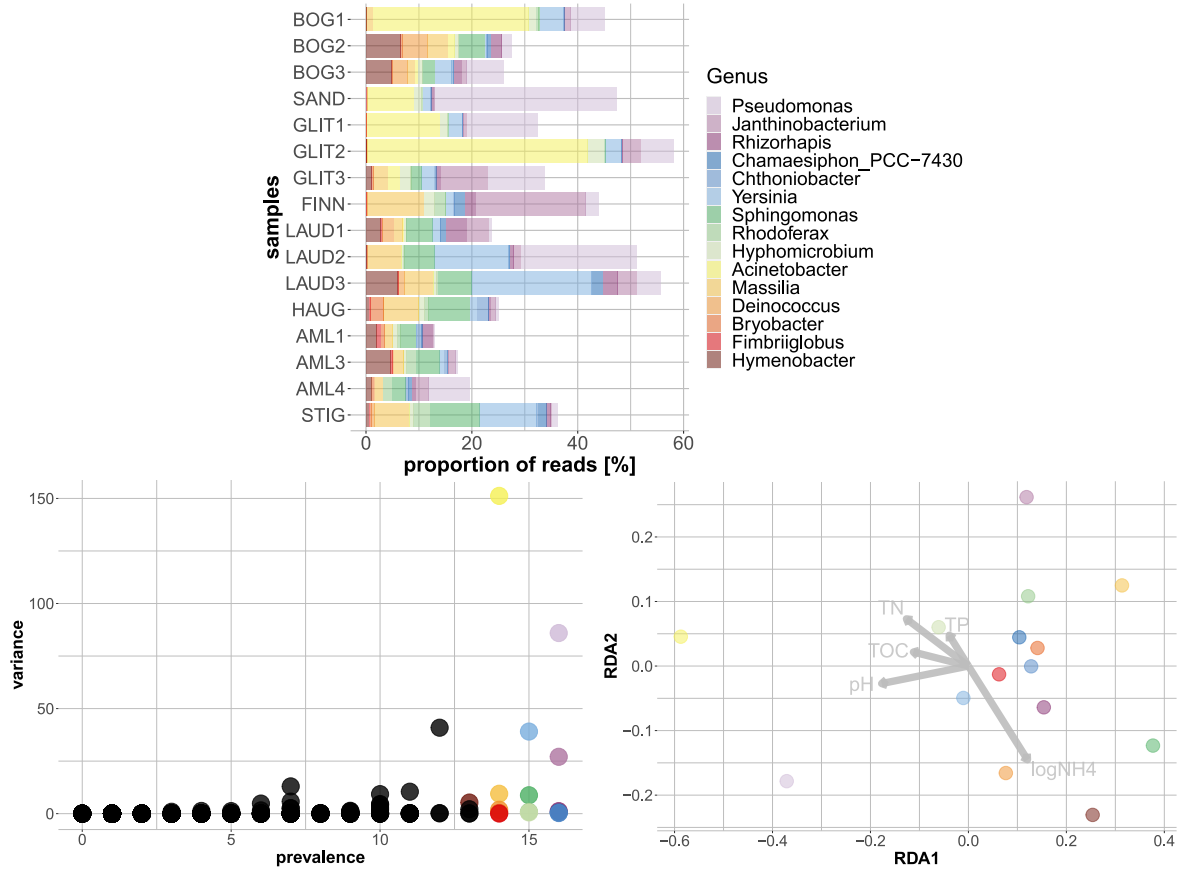


Figure 7. Bacterial genera from the 16S sequencing data showing the highest prevalence and variance. The RDA shows the relationships among environmental variables and specific genera among the most prevalent and abundant. Samples named BOG refer to the Lysaker river, SAND to Iselva river and GLIT to Lomma river, all belonging to the Oslo region. In the Agder region instead we had the Finnsåna river (FINN), Mandal river (LAUD), Haugedøl river (HAUG), Nidelva river (AML) and Stigvasselva river (STIG).

Paper II

Paper II focused on the effect of hydrological alterations due to the impact of hydropower and dams on Norwegian rivers and the structure and functioning of biofilm microbial community. This study was carried out during two sampling campaigns, the first being in June 2019 and the second in August 2019. Ten river systems were selected from one of the study areas used also for **paper I**, the Agder region in south western Norway, well known to be massively exploited for hydroelectric energy production. The selected rivers where somewhat comparable in size and they were all impacted by hydropower or dam developments. The design of our study was the control-impact type, with the dam/hydropower development being the source of impact. Thus, for each river, a paired design was used with one site upstream the dam or powerplant, representing a free-

flowing river section and a second site downstream of the dam. Both sites had comparable ecological conditions. In total, there were 20 sampling sites where biofilm was sampled for DNA sequencing, water samples were taken for chemical analyses, the fluorescence of the autotrophic components was measured, and the decomposition rates of leaf litter quantified. In addition, one more sample of water to analyze the chemistry was taken at the outlet of each powerplant above the mixing zone and after it had been through the turbines, to have an overview of the impact of the hydropower on the water chemistry.

Ceramic tiles were used (**Fig. 8**) to grow the biofilms within the river systems, standardizing in this way the differences due to the geology of rocky bottom. DNA from the biofilms grown on the tiles were used for 16S rRNA gene metabarcoding and for the analysis of the prokaryotic community composition as for **paper I**.

Figure 8. Ceramic tiles were our standardized substrate for biofilm growth in the control and impacted river reaches (Photos: Lorenzo Pin).

In addition, litter bags were employed, with coarse and fine mesh to separate the effects of the microbial communities from the whole benthic community. As a proxy for organic matter, cloths stripes made of 70% cotton and 30% cellulose (Wettex, Vileda GmbH) were used (**Fig. 9**).

Figure 9. Wettex stripes were used as proxy to measure the organic matter breakdown in the impacted and unimpacted reaches. After the exposure period the dry weights were measured and then the samples were burned to calculate the Ash Free Dry Weight.

Dry weights and AFDM were measured for each of the replicate, and the data are reported as AFDM loss percentage using the following formula:

$$\left\{\frac{[AFDM(t0) - AFDM(t)]}{AFDM(t0)}\right\} * 100$$

Where AFDM(t0) stands as the initial airdried weight times the correction factor D, which was calculated from a set of Wettex stripes which was not put into the rivers but was weighted to measure the average AFDM/initial-airdried-weight. AFDM(t) is the value of AFDM registered for a replicate at the end of the exposure period.

The tiles and the litter bags were placed in two sampling sites for each river: one site, representing the control, was upstream the dam and the second site was downstream and represented the impacted site. This control-impact design aimed at identifying generalized patterns for the microbial biodiversity and for the breakdown of organic matter by the biological communities between the control and impacted sites within each river and among the 10 rivers analyzed.

No significant impact from the hydropower or dam was detected for the environmental variables. Rather than showing patterns related to the flow alterations due to the impoundment environmental conditions were specific to each river. Light availability was the only environmental variable showing a significant difference when looking at upstream downstream gradients. Downstream reaches had on average a lower light intensity compared to the upstream sites (**Fig. 10**).

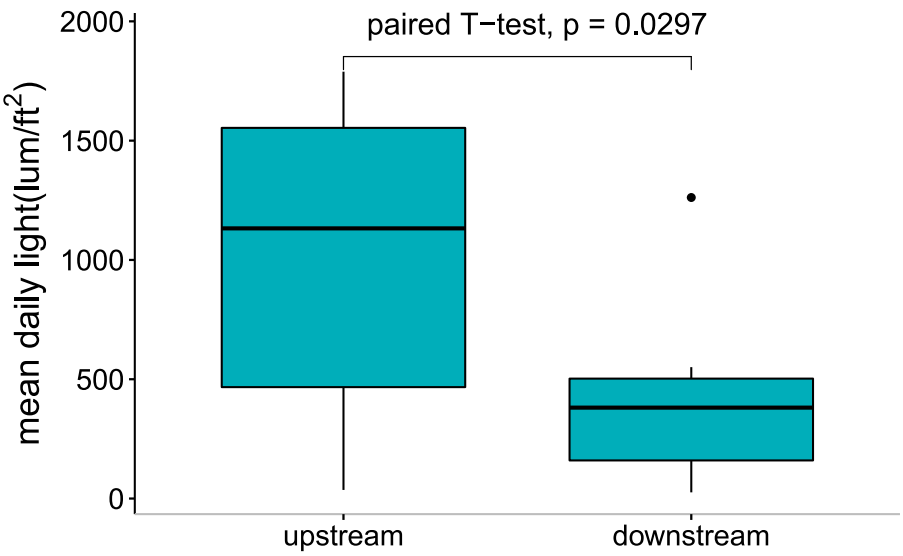


Figure 10. Mean daily light intensity for the period of exposure showed significantly lower values in the downstream (n=10) reaches compared to the upstream (n=10).

Significant differences were detected in the decomposition of organic matter, being higher in the downstream reaches than in the upstream, in particular for coarse mesh bags. Some fine mesh bags were lost and thus, the statistical power for the comparison between upstream and downstream sites was affected. However, a general tendency for higher decomposition in the downstream reaches was observed also for the fine mesh bags (**Fig. 11**).

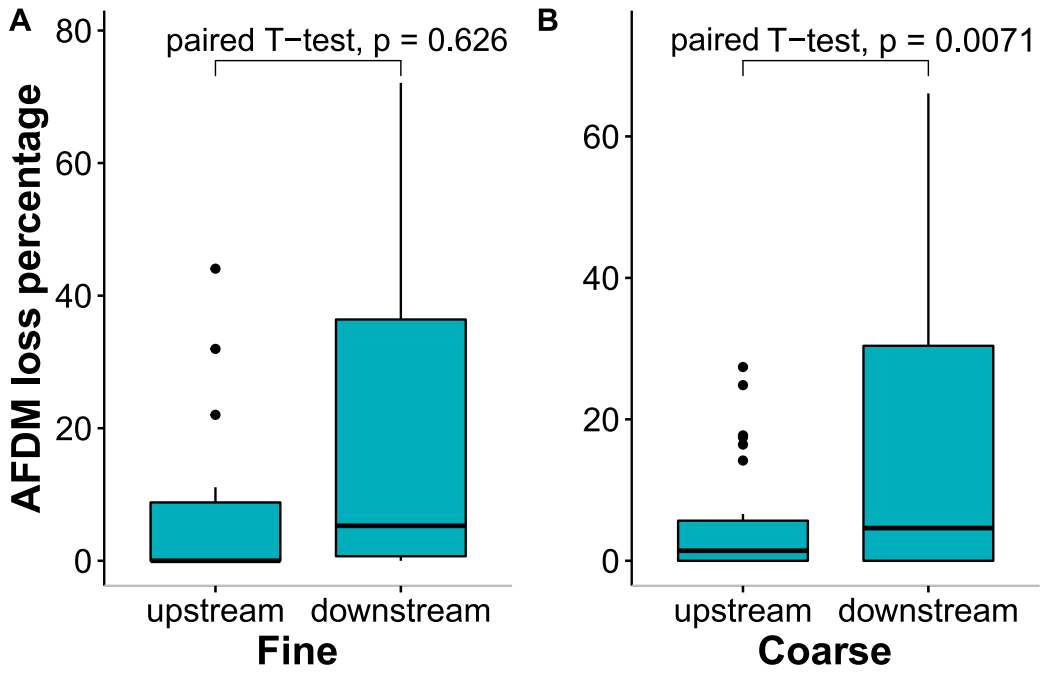


Figure 11. Significant differences were found for the AFDM loss percentage in the downstream reaches compared to the upstream for the coarse mesh bags (df = 29). Although no significant differences were found for coarse mesh bags (df = 8, due to missing bags).

The 16S sequencing data on the microbial community structure, revealed that the main factor driving the prokaryotic assemblages was not ascribed to the hydropower or dam impact, but to the source catchment, which overall masked the more subtle effects of the flow fragmentation. The upstream and downstream communities shared 20% of the Amplicon Sequence Variants (ASVs), which makes up to a consistent share of the prokaryotic community, but 37% belonged only to upstream and 43% to downstream, meaning that the majority of ASVs were specific to each river reach (**Fig. 12**). The autotrophic component did not show significant changes when looking at upstream and downstream gradient, neither from the benthotorch results nor for the cyanobacterial ASVs relative abundances.

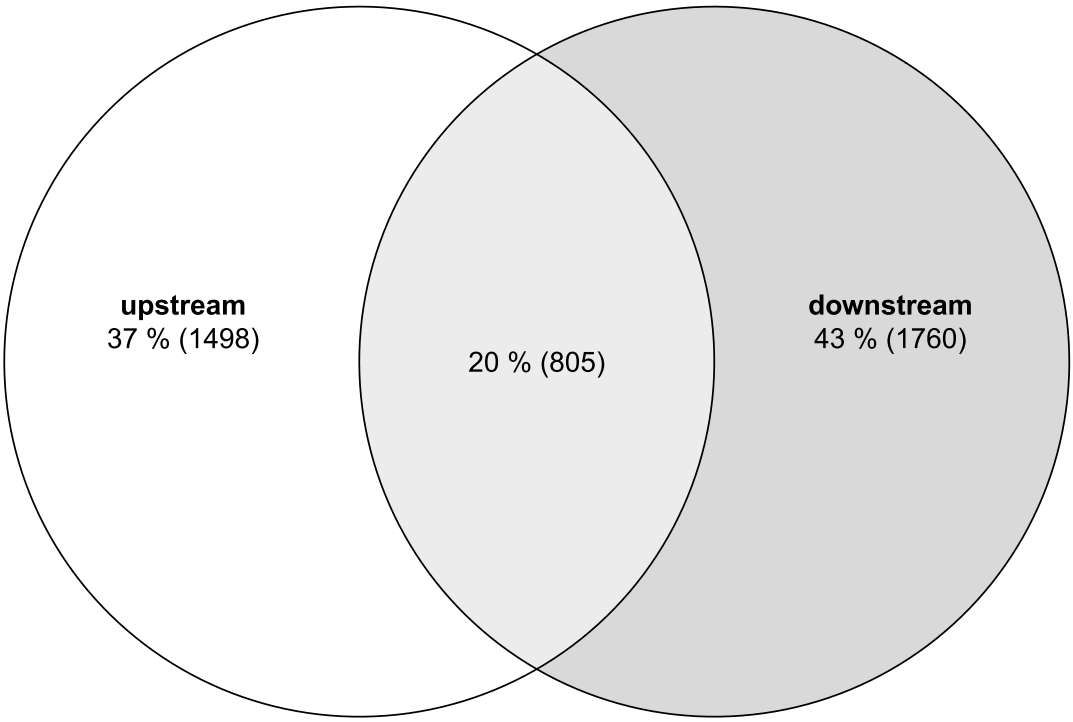


Figure 12. Venn diagram showing the different fractions of Amplicon Sequence Variants shared by both the upstream and downstream sites and the share belonging only to each of the river reach.

Paper III

Paper III was based on a mesocosm experiment in 16 stainless steel flumes, performed at the EcoLaboratory facility at the University of Birmingham. Four different water flow scenarios were simulated and the response of the biological communities in two different habitat types within each flume was analyzed (**Fig. 13**).

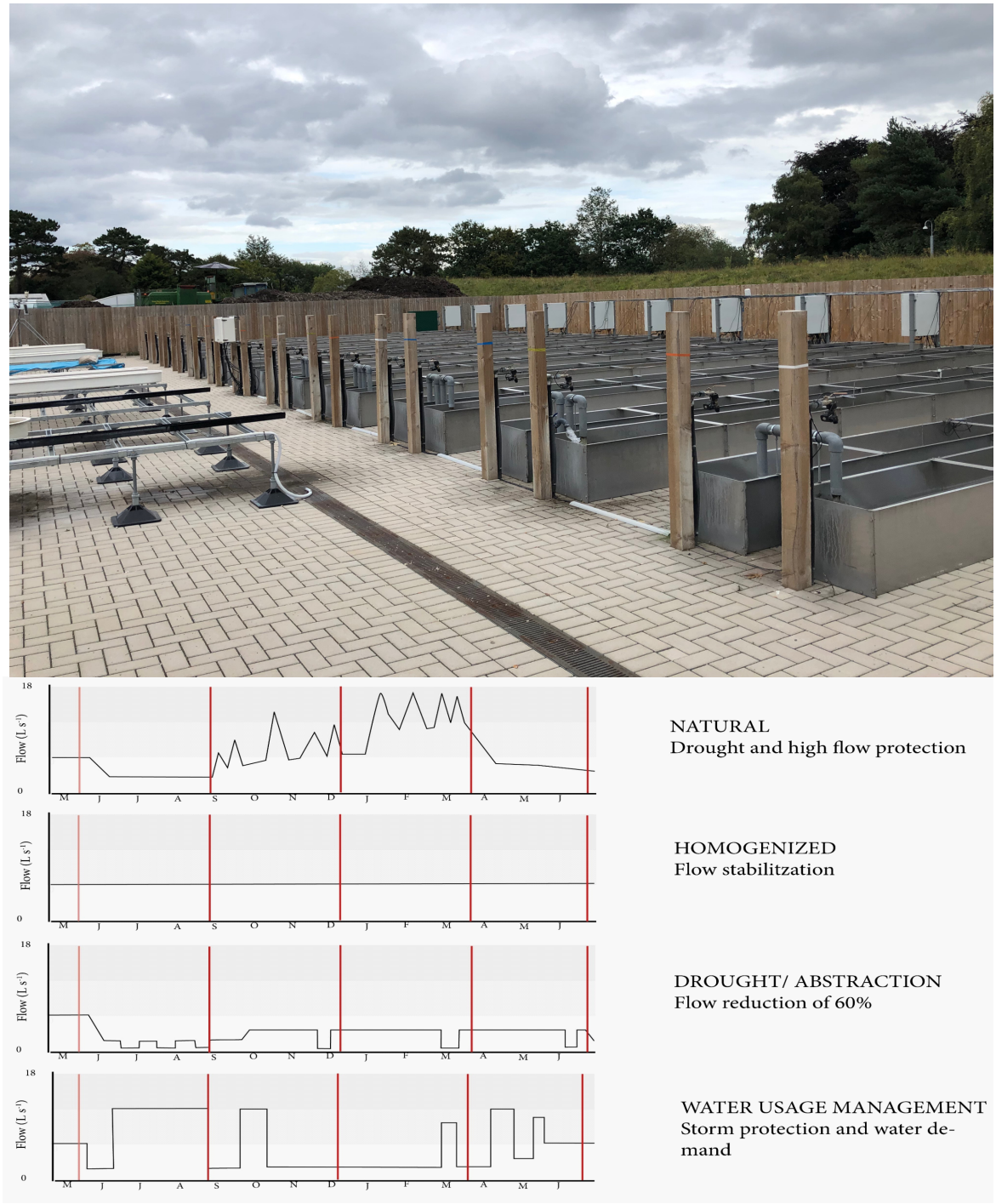


Figure 13. In the picture above, we can observe the 16 stainless steel flumes which were deployed for the experiment on the response of microbial community structure and functioning to different flow regimes at EcoLaboratory facility at the University of Birmingham, UK (Photo by Raquel Arias Font). Below the picture the schematic overview of the 4 different flow treatments that were implemented.

A generalized randomized complete block design was implemented with two blocks and 4 treatments with 4 replicates. The flumes were 12 m long and 0.5 m wide, and slightly inclined to promote the water flow downstream. Each of them was divided into 3 sections, and each section

was divided in two different habitats, a riffle and a pool. The discharge in the flumes was implemented by automatized taps which could provide up to 18 L/s.

The bottom of the flumes was covered with evenly sorted gravel and two species of macrophytes were planted in the habitats, watercress *(Nasturtium officinale*) in the riffles and water milfoil (*Myriophyllum* sp.) in the pools. Organic matter and macroinvertebrates were sampled from a nearby stream (Bourn brook) and added in equal quantities to the flumes to establish the community.

The four different flow treatments are representative for rivers impacted by human management. The first treatment consisted in a simulation of hydrographs from rivers affected by management for irrigation purposes and storm protection. These conditions are characterized by higher flows during the dry season due to irrigation and lower flows during the rainy season, when the water is retained within the reservoirs. The second treatment simulated a hydrograph of a river subjected to water abstraction, with around 60% less water than a natural free flowing river. This treatment was regarded as drought treatment, as during the dry season, extreme drought events occurred. The third type of flow regime implemented as a treatment in the flumes, simulated a homogenized hydrograph, with a constant discharge during each season. This is a condition common to many rivers that are affected by dams where no natural variability is implemented in the downstream outlets. The fourth treatment was the control and simulated the variability of a natural flow regime, with high flows during rainy season and during the ice-melt season, low flows during summer and winter with occasional droughts. The flumes were kept running since October 2018, and our experiment was performed from November 2019 to February 2020.

As for the experiment in **paper II**, a standardized substrate was selected to study the organic matter breakdown and also as substrate for biofilm growth. Fifteen coarse mesh bags and three fine mesh bags (200 μm) containing (on average 2.26 ± 0.09 g) cloth stripes (Wettex, Vileda GmbH) were placed in both pools and riffles in the second section of each flume. Ten replicates (five from each habitat) of the coarse mesh bags were collected at three different sampling times:

after 29, 64 and 100 days of exposure. All the fine mesh bags were collected at the third sampling time, after 100 days, to ensure enough decomposition. Decomposition data were reported with the same technique for the experiment in **paper II**. The data for organic matter breakdown were analyzed by linear mixed effects models to test the effect of sampling time, treatment and habitat. The Wettex replicates from each habitat were gently squeezed and the biofilm removed and pooled to obtain one sample from each habitat in each flume. For the coarse mesh bags, 32 samples (16 riffles and 16 pools) were collected for each sampling time. For fine mesh bags 32 samples in total were collected.

The slurry obtained for each sample was then filtered through a 0.2 µm polycarbonate membranes (47 mm diameter, Nuclepore) by gentle vacuum (< 0.2 bar). The filters were then stored at -10 °C. DNA extraction was performed from the filters and 16S rRNA gene sequencing performed to obtain metabarcoding data for the prokaryotic communities from the different treatments across the three sampling times for the two habitats.

The results from the organic matter breakdown showed significant effects due different flow treatments. Overall the treatment having the clearest impact on the organic matter decomposition was the one implementing a flow reduction of 60%, which had always a significantly higher decomposition compared to the other treatments which on the other hand did not show significant differences among each other.

The model best explaining the organic matter breakdown variation for the coarse mesh bags included interactions between the treatments and habitat and an additive effect of sampling time (AIC = 1096.2). For the fine mesh bags, the best model included a three-way interaction between the treatment, habitat and block (AIC = 208.5). The treatment-habitat interaction was stronger compared to the treatment-block interaction. Overall, the treatment effect was the strongest (**Fig. 14**).

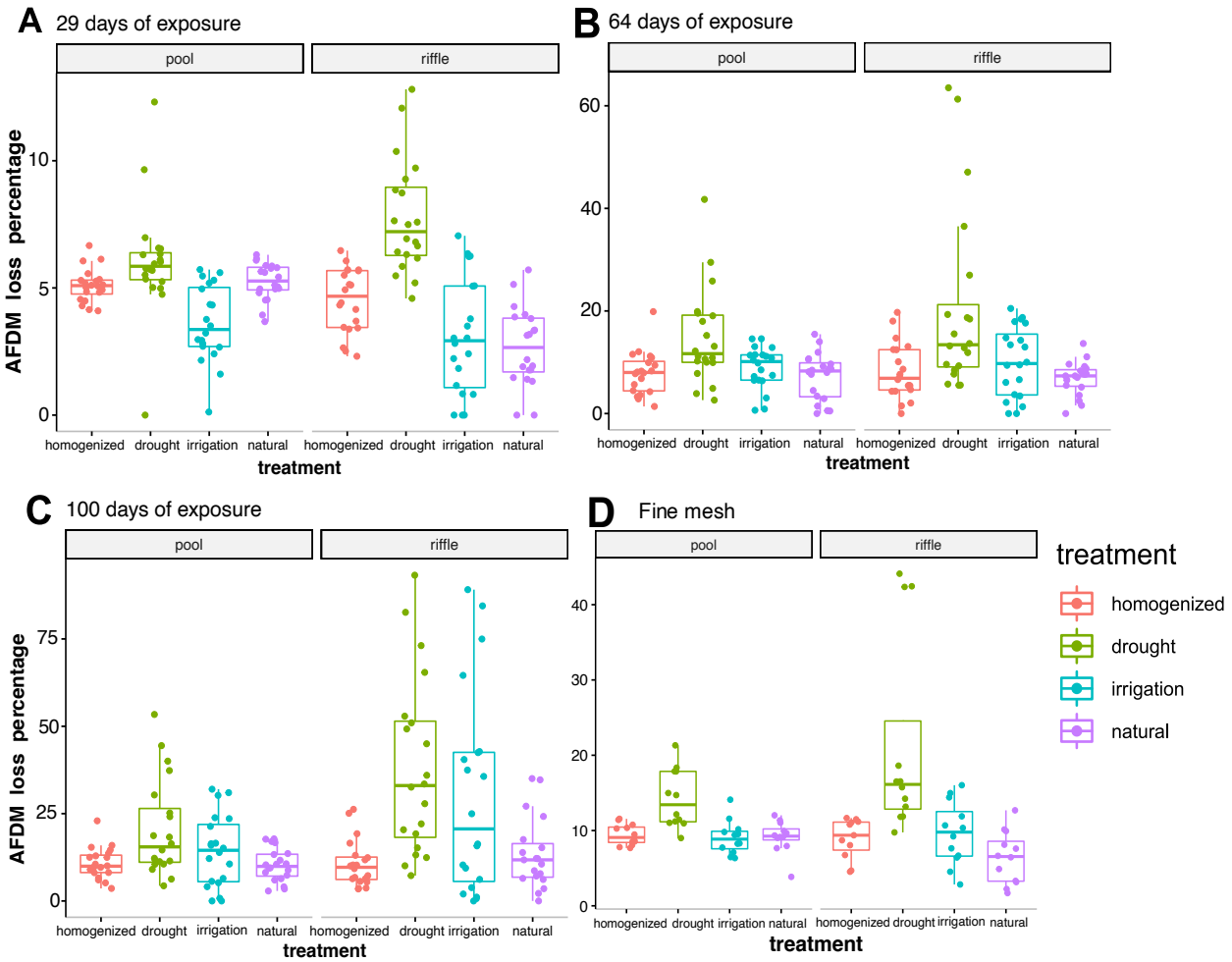


Figure 14. Organic matter decomposition calculated as Ash Free Dry Mass Loss percentage for the coarse mesh bags after A) 29; B) 64; C) 100 days of exposure. D) Fine mesh results after 100 days of exposure.

The prokaryotic community composition was also significantly impacted by the different flow regimes. Beta diversity was significantly affected by the flow treatments for both the fine and the coarse mesh bags, as shown by pairwise PERMANOVAs performed on the Bray-Curtis matrices (**Table 1**).

Table 1 Significant differences were found when comparing the Bray-Curtis dissimilarity matrices for the different treatments. The results showed how each treatment had a specific biofilm community composition, for both the coarse and fine mesh bags. Sampling times also showed a significant effect on the community composition for the coarse mesh bags, which significantly changed over the three different sampling times.

Pairwise Adonis on Bray-Curtis dissimilarity matrix

Fine mesh

pairs	Df	SumsOfSqs	F.Model	R2	p.value	p.adjusted	
homogenized vs drought	1	0.467	2.168	0.134	0.001	0.006	*
homogenized vs irrigation	1	0.540	2.785	0.166	0.002	0.006	*
homogenized vs natural	1	0.530	2.531	0.153	0.001	0.006	*
drought vs irrigation	1	0.591	2.712	0.162	0.001	0.006	*
drought vs natural	1	0.599	2.567	0.155	0.001	0.006	*
irrigation vs natural	1	0.340	1.609	0.103	0.017	0.017	.

Coarse mesh (T1)

pairs	Df	SumsOfSqs	F.Model	R2	p.value	p.adjusted	
homogenized vs drought	1	0.432	1.357	0.088	0.086	0.094	
homogenized vs irrigation	1	0.444	1.438	0.093	0.047	0.094	
homogenized vs natural	1	0.568	2.400	0.146	0.001	0.006	*
drought vs irrigation	1	0.571	1.852	0.117	0.001	0.006	*
drought vs natural	1	0.701	2.972	0.175	0.001	0.006	*
irrigation vs natural	1	0.469	2.068	0.129	0.001	0.006	*

Coarse mesh (T2)

pairs	Df	SumsOfSqs	F.Model	R2	p.value	p.adjusted	
homogenized vs drought	1	0.542	2.167	0.134	0.002	0.010	*
homogenized vs irrigation	1	0.481	1.956	0.123	0.008	0.016	.
homogenized vs natural	1	0.437	1.829	0.116	0.004	0.012	.
drought vs irrigation	1	0.652	2.558	0.154	0.001	0.006	*
drought vs natural	1	0.571	2.304	0.141	0.002	0.010	*
irrigation vs natural	1	0.352	1.447	0.094	0.053	0.053	

Coarse mesh (T3)

pairs	Df	SumsOfSqs	F.Model	R2	p.value	p.adjusted	
homogenized vs drought	1	0.480	2.172	0.134	0.001	0.006	*
homogenized vs irrigation	1	0.448	1.934	0.121	0.008	0.008	*
homogenized vs natural	1	0.485	2.629	0.158	0.001	0.006	*
drought vs irrigation	1	0.584	2.444	0.149	0.001	0.006	*
drought vs natural	1	0.712	3.714	0.210	0.001	0.006	*
irrigation vs natural	1	0.401	1.978	0.124	0.001	0.006	*

Coarse mesh sampling times

pairs	Df	SumsOfSqs	F.Model	R2	p.value	p.adjusted	
t1 vs t2	1	0.672	2.315	0.036	0.001	0.003	*
t1 vs t3	1	0.906	3.250	0.050	0.001	0.003	*
t2 vs t3	1	0.473	1.740	0.027	0.002	0.003	*

Signif. codes: '*' p <.05; '**' p <.01; '' p <.001***

However, when looking at alpha diversity measures (Chao1 index) and evenness (Pielu's index), no significant effect was found for the different flow treatments. These results highlight a clear impact of flow alterations implemented for the different flumes on the prokaryotic species turnover, while the number and the dominance relationships among the species for each treatment seemed to be relatively stable. Significant differences were also found among the prokaryotic communities from different sampling times, reflecting the ecological succession and the development of the community from an initial colonizing stage to a more mature biofilm.

The treatment showing the most evident effects on both the microbial community, as for the organic matter degradation, was the drought treatment, which was always significantly different from the other flow treatments, as also shown in the RDA (**Fig. 15**).

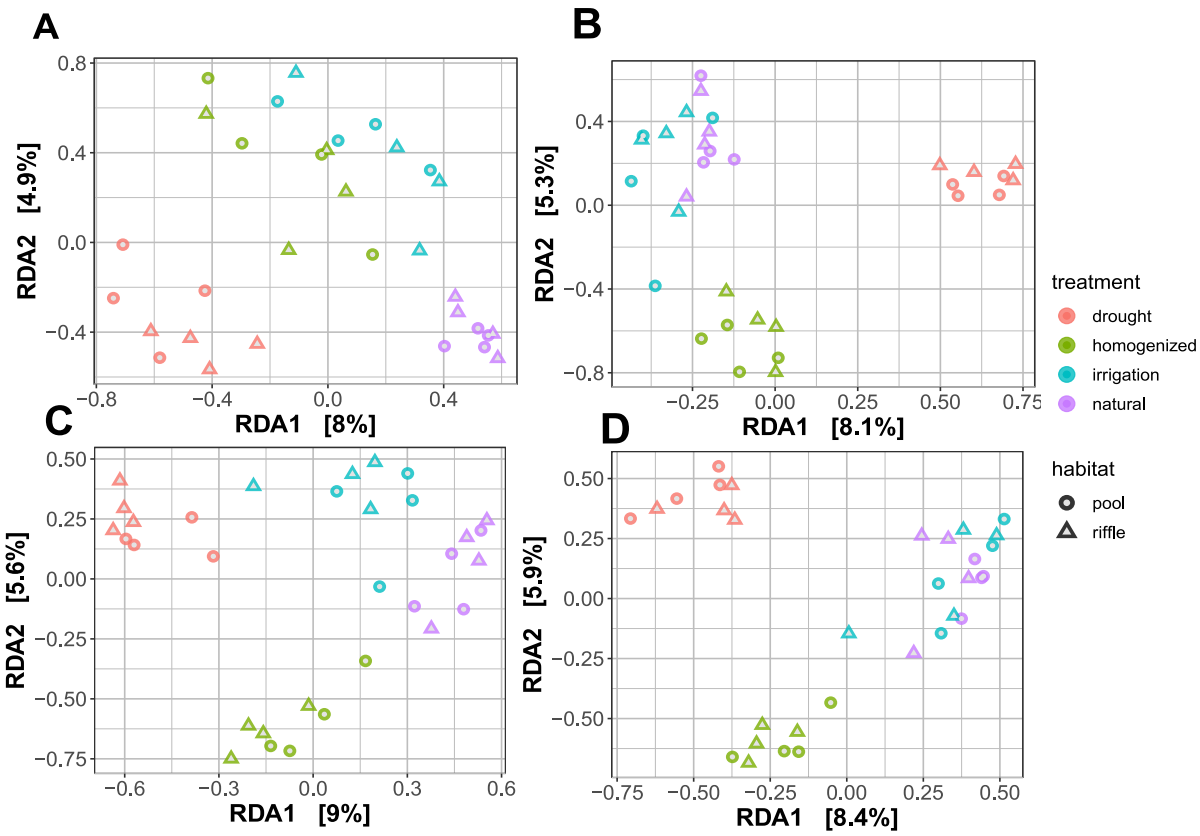


Figure 15. Visualization of the prokaryotic community structure by a treatment-constrained distance-based Redundancy Analysis for the coarse mesh bags after A) 29; B) 64; C) 100 days of exposure. D) Prokaryotic community structure for the fine mesh bags in the different treatments after 100 days of exposure.

Paper IV

In paper IV we explored different methodologies to find suitable biological indicators for water quality and the overall ecological status of the river system. In this study we used existing datasets of environmental and biological variables for the Danube River system collected during the third Danube Joint Survey in 2013 organized by the International Commission for the Protection of the Danube River (ICPDR; http://www.icpdr.org/wq-db/). The dataset considered for this study comprised the left, right and middle transects for the Danube River, with in total 160 samples. The

planktonic prokaryotic community was analyzed by using 16S rRNA metabarcoding and only the bacterial taxa were used for the analyses.

From the Amplicon Sequence Variants (ASVs) table we created three datasets at phylum, class and genus level. Following the methodology developed in **paper I**, we calculated the frequency of occurrence of the *taxa* for each of the three datasets. After that, the coefficient of variance for the relative abundances of each specific *taxa* was calculated (standard deviation of taxon abundance divided by the mean). The results from the two analyses were then plotted in a scatter plot and the taxa with highest variance and prevalence were highlighted as suitable biological indicators given their cosmopolitan nature and their high variability which could be related to shifts in specific environmental drivers. To detect the environmental feature driving the abundances of each specific taxa for the three different datasets a distance-based Redundancy Analysis was performed, including the environmental parameters which are usually good indicators for water quality (Chlorophyll-a, total Nitrogen, Dissolved Organic Carbon, Water Temperature, pH and Conductivity, Total Coliformes, *E. coli*, Total bacteria, Human related Bacteria). By using a step forward analysis, the best explanatory parameters for the microbial community structure were selected and plotted to visualize relationships between the single taxa and the environmental drivers. Among the most informative taxa at phylum level, Actinobacteriota, Bacteroidota, Verrucomicrobiota, Proteobacteriota, Cyanobacteriota and Planctomycetota were the top six. The first four taxa were also among the most abundant overall in the Danube, with the Actinobacteriota being dominant in the community. The RDA results showed how Actinobacteriota were closely related to low levels of pH, Chlorophyll-a and overall low nutrient contents, which were typical of the lower reaches of the Danube where the dilution was reducing the concentration of nutrients and organic matter related compounds. On the other hand, Bacteroidota were closely related to the upstream river reaches, characterized by higher nutrients concentrations where the inputs of allochthonous organic matter from the forested areas were significantly higher than in the lower reaches. Verrucomicrobiota showed strong positive

relationship with chlorophyll-a concentrations. At class level, similar behaviors in the microbial community have been observed (**Fig. 16**).

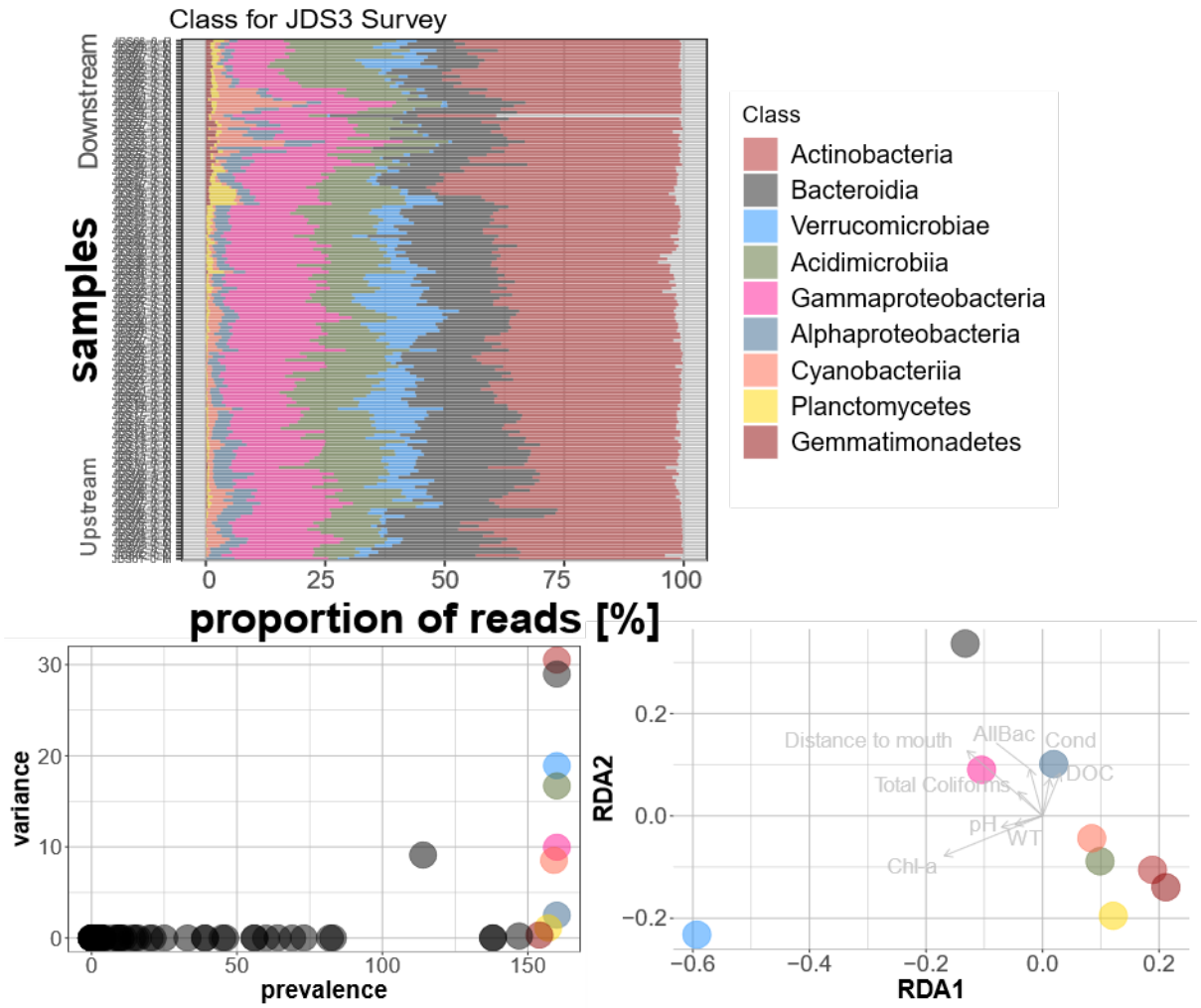


Figure 16. Prevalence/Variance analysis on the prokaryotic taxa at class level for the Danube river. The scatter plot highlights the potential bioindicators based on their occurrence and variance values. The relative abundances for the 10 most relevant taxa are highlighted in the barplot. The RDA shows the relationships among the potential bioindicators and selected environmental variables relating to water quality.

Actinobacteria are in fact the class with highest prevalence and variance values overall, showing a positive relationship with low nutrient levels and low chlorophyll-a content, this group increases in abundance from the source to the lower reaches of the Danube River. Opposite pattern is found instead in the Bacteroidia, for which the abundances showed positive relationships with high concentrations of dissolved organic carbon, conductivity. Overall, their abundances were highest in the uppermost river reaches, decreasing significantly while reaching the mouth of the Danube.

The strong positive relationship between Verrucomicrobiae and chlorophyll-a concentration confirms the results seen at phylum level, this taxon seems to be also positively influenced by higher pH and water temperature. Among the genera, the dominant one in terms of prevalence and variance was the *hgcl_clade* of the Actinobacteria class, confirming the general patterns of the prokaryotic community observed at class and phylum levels.

This taxon seemed to be strongly related to low pH levels, low conductivity and overall, its abundances were higher the closer to the river mouth. On the opposite side of the hyperdimensional space we found the *Flavobacterium,* taxon belonging to the class of Bacteroidia, confirming the opposite behavior and environmental preferences for these two taxonomic groups. Overall, *Flavobacterium* were more abundant where pH, conductivity and nutrient concentrations (in particular Nitrogen) were highest, that is in the river reaches farer from the river mouth, in the upper section of the Danube. A similar pattern was observed for another genus, *Sediminibacterium2*, another taxon belonging to the Bacteroidia class, particularly abundant in the upper reaches of the Danube, where the loads of particulate organic matter, nutrients and pH were highest.

The second methodology used to detect prokaryotic bioindicators for water quality focused on the implementation of different machine learning algorithms. An unsupervised random forest (RF) model was used with the ASVs table for the Danube to detect the underlying patterns in the microbial community structure and reveal the ASVs with a positive contribution to the clustering of the sampling sites. This step served as a filter to retain only the informative ASVs for the following analyses. For the optimization of the parameters (mtry and ntree) of the unsupervised model model we run a grid search using a supervised random forest model, with chlorophyll-a concentration as response variable and ASVs as predictors with a 0.8/0.2 dataset split for training and testing sets. Only the ASVs yielding an increase in the percentages of the mean squared error (MSE) were used in the following analyses. A grid search was then run using XGboost to search for the best number of steps for the transformation and lag of each of the selected ASV. Each

ASVs was added (with their best transformation and lag values) to the initial design matrix (containing sample ID, transect code and distance to mouth) only when the R squared of the model increased, otherwise it was discarded. After a first screening of all ASVs, the process of screening continues until there are no more changes to the ASVs selection. As we were running the models using XGboost we therefore needed to optimize the hyperparameters of the best model, running several iterations to find the best combination. For this model the training and test fractions are 0.85 and 0.15 as some of the water quality classes were occurring only after 550 km from the river mouth, so past the 80% of the dataset. The best models for water quality were evaluated looking at the prediction accuracy and only the ASVs combination with 100% accuracy were selected, fort he chlorophyll-a prediction instead, we selected the the model yielding the best R squared overall.

The list of ASVs detected in the best models for both water quality and chlorophyll-a were then used to build a network analysis, to show patterns of co-occurrence and co-exclusion of the identified ASVs. The results highlighted a set of ASVs which were most informative for water quality predictions, in particular *Rhodoluna* and *Flavobacterium*, belonging respectively to the Actinobacteriota phylum and the Bacteroidota, were regarded as the most informative as they occurred more often than the other taxa in the best predictive models. Although these two were the main taxa, there were several other ASVs with which had complementary shares of predictive information (**Fig. 17**).

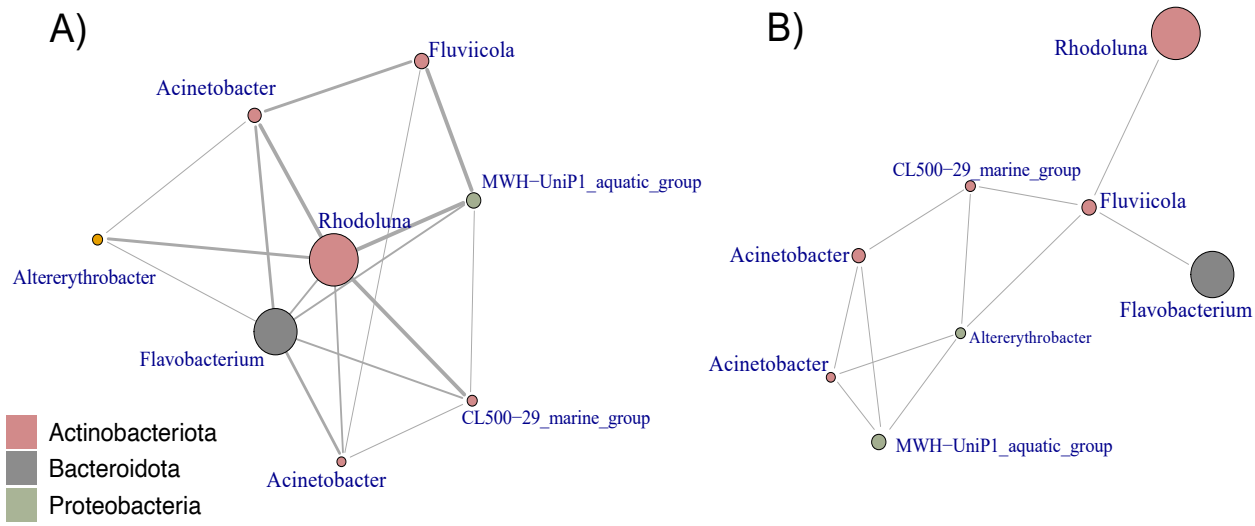


Figure 17. Network analysis results for the most frequent ASVs returned by the spatio-temporal approach yielding perfectly accurate water quality classification. Colors represent the bacterial phylum each taxon belongs to. A) Co-occurrence network, the dots' dimension represents the number of occurrences of each taxon in the best models, while the thickness of the links represents the number of co-occurrences for ASV pairs. B) Co-exclusion network. Links represent taxa that never occurred together in a same model.

The results from the machine learning models allowed us to identify combinations of ASVs which could be used in the prediction of the water quality of a selected study sites by looking at the data from upstream sites, due to the organisms' mobility within the river.

Discussion

Molecular methodologies for the analysis of prokaryotic community structure in Norway (Paper I)

Paper I served as a first screening of the prokaryotic communities living in epilithic biofilms in Norwegian rivers impacted by anthropogenic and natural stressors, which was still a black box given the lack of studies focusing on prokaryotes from this country. The novelty of the study concerned the joint use of two different molecular techniques to characterize the prokaryotic assemblages at different taxonomic levels. The combination of a quantitative method such as CARD-FISH and a semi-quantitative technique like 16S rRNA gene amplicon sequencing, allowed to obtain a comprehensive overview of the diversity and abundance of the prokaryotic

community composition of epilithic biofilms, which are still poorly understood in the context of riverine ecosystems. This study highlighted the importance of quantitative methodologies like CARD-FISH to analyze the absolute prokaryotic abundances and activity of specific taxa, detecting small scale variation in the microbial communities which are not identified by sequencing approaches (Freixa et al., 2016; Bakenhus et al., 2019; Fazi et al., 2020).

Peaks in bacterial and archaeal abundances revealed a great variability among the different sampling sites, even within the same river. This precision could allow the detection of point sources of pollution within river systems. Comparisons among the abundances of the dominating bacterial groups detected by CARD-FISH, highlighted potential discrepancies in the community dynamics due to different nutrient loads and the overall trophic status of river systems. However, the high variability shown by the CARD-FISH results at fine spatial scale resolution, might mask the effect of the main environmental drivers at community level. No relationships were found among the prokaryotic absolute abundances and specific environmental variables. Cell numbers might be affected by microscale ecological features that were not measured, such as interactions with other biological or physico-chemical components. Variability might also be caused by random events such technical or sampling biases.

Metabarcoding seemed to be more reliable when analyzing communities across large spatial scales. The huge amount of data and the fine taxonomic resolution provided by metabarcoding enables the exploration of large-scale patterns among environmental drivers and prokaryotes at different taxonomic levels at the same time (Ligi et al., 2014), which is not possible when using hybridization in situ methodologies (Bouvier & del Giorgio, 2003; Corte et al., 2013). The biofilm communities across all the samples were overall dominated by the phylum of Proteobacteria, confirming the findings from previous studies on freshwater bacteria (Battin et al., 2016; Besemer et al., 2012; Wilhelm et al., 2013). Alphaproteobacteria and Gammaproteobacteria were the most abundant classes, as confirmed by both, CARD-FISH and sequencing results. Although, for the CARD-FISH results, the Alphaproteobacteria were the most abundant overall, while the relative

abundances from the sequencing results showed a dominance of the Gammaproteobacteria. This contrasting finding might be related to updates in the classification in the "silva_nr_v132_train_set.fa" (Callahan, 2018), where Betaproteobacteria, formerly a class of the Proteobacteria, are now an order under the class of Gammaproteobacteria, whereas for the CARD-FISH analysis two different probes were used to identify these taxa, treating them as separate entities. Therefore, the abundances of the two groups were joined at a later stage to align with the sequencing results.

Among the Gammaproteobacteria genera identified by the metabarcoding analysis, the most represented *taxa* were *Pseudomonas*, *Acinetobacter* and *Yersinia*, all of them being important players in the biofilm formation process (Williams et al., 1996; Lee et al., 2008). From the Alphaproteobacteria, the most represented genus was *Sphingomonas*, crucial in the structural composition of biofilms (Johnsen et al., 2000). Other represented genera were *Massilia* and *Janthiniobacterium*, both belonging to the order of *Burkholderiales*, which is regarded as one of the important bacterial functional groups in freshwater environments as they host taxa capable of ammonium oxidation, crucial for the cycling of nitrogen (Zhang et al., 2012; Gołębiewski et al., 2017). The main driver for the microbial community structure seemed to be the sharp gradient in the pH of the river systems analyzed, which decreases considerably from the Oslo region to Agder region, due to the different bedrock and historical acid rain deposition. Higher pH are known to positively influence members of the Alphaproteobacteria (Bragina et al., 2012), which were the dominant community in the Agder region, according to the CARD-FISH results. Although a significant effect of pH on prokaryotic communities was detected, it was impossible to disentangle the effects of other environmental variables from the strong regional pattern, which was the main variable explaining the difference among the composition of the prokaryotic communities.

The comparison of the two techniques employed in this study provide different insights into the complexity of the prokaryotic community structure of epilithic biofilms in Norwegian river

systems. We consider sequencing to be a more effective tool for large spatial scales analyses, as the results we obtained were closely related to the environmental gradients and enabled us to identify specific taxa which could be closely related to specific environmental drivers. On the other hand, CARD-FISH allows to quantitative analyze the actual abundances of specific taxa to detect their relationship with specific variables. The parallel use of these two methodologies would be useful to detect taxa suited as bioindicators for pollution, potential pathogens or eutrophication.

Response of prokaryotic community structure and functioning to modified flow regimes (Paper II-III)

Hydropower impact in Norwegian river systems

In **paper II** the effect of river fragmentation by hydropower and dams and the resulting flow alterations, were analyzed more in depth in relation to the biofilm prokaryotic community structure and functioning. The study was based on a control impact design, where upstream unimpacted river reaches and downstream impacted river reaches were compared across 10 different river systems, all impacted by hydropower developments.

The results from the organic matter decomposition from the coarse mesh bags showed a significant difference between impacted and unimpacted river reaches, with the former ones having higher values of AFDM loss percentage. The fine mesh bags showed a similar pattern, although, the lack of replicates (some of the bags were flushed away) made it not possible to obtain statistical significance when performing the pairwise comparisons among the functional data for the single rivers. Moreover, the sequencing data of the prokaryotic biofilm assemblages did not show significant differences when comparing upstream and downstream sites. Both, alpha diversity measures and species turnover were tested to assess if a generalized effect of the river impoundment was mirrored in the prokaryotic community. This was not the case and upstream and downstream sites were in general similar in terms of prokaryotic taxonomy and relative

abundances. The biofilms in this study were overall dominated by Proteobacteria and Bacteroidota. Cyanobacteria, which were one of the target taxa and expected to be significantly higher in relative abundances in the downstream communities due to the reduced flow variability, were among the most abundant bacterial taxa. When analyzing the autotrophic community more in depth, the ASVs belonging to the Cyanobacteria did not show significant differences in upstream and downstream patterns. These results were also confirmed by the benthotorch measurements on the different chlorophyll spectra related to the autotrophic components of epilithic biofilms (green algae, cyanobacteria and diatoms). The sequencing results were reflected also by the water chemistry parameters, which were overall similar for both upstream and downstream sites, confirming the close relationship among prokaryotic communities and environmental variables found in **paper I.**

The significant difference in organic matter decomposition between impacted and unimpacted sites might be related to an increase in the activity of the heterotrophic community downstream the impoundments. The variability of the downstream results was also significantly higher than the upstream sites, this finding may reflect differences in the downstream flow management by the single hydropower companies and river specific biological responses to the flow alteration. The results from the coarse mesh bags report the effect of the whole benthic community, including both microbial and macroinvertebrate communities. The latter are crucial players in the decomposition of recalcitrant organic matter in river ecosystems (Gessner et al., 1999), and their feeding activity might also explain the higher degree of decomposition in the coarse mesh bags compared to the fine mesh ones.

Increased biofilm activity in the downstream reaches may be due to the tight link between the activity of detritivores and the conditioning of organic matter carried out by the microbial communities. The early stage colonization of organic matter by microbes renders the substrate more palatable and nourishing for the macroinvertebrates (Foucreau et al., 2013). The increased decomposition in the downstream sites recorded for the fine mesh bags is also directly linked to

the microbial activity, and although this discrepancy was not significant, a general tendency of higher decomposition processes mediated by microorganisms was observed in the impacted sites. Stable low flow regimes are a common feature of hydropower impacted rivers in Norway (Alfredsen et al., 2012) and, flow stability is recognized to be a driver for increase in autotrophic activity and production of labile organic matter, which can stimulate the heterotrophic component of the biofilms (Aristi et al., 2014).

Light intensity was the main environmental parameter which significantly differed between impacted and unimpacted reaches, with higher values for the latter. The lack of light in nutrient poor environments is related to enhanced heterotrophic activity and breakdown of organic matter due to increased autotrophic production (Howard-Parker et al., 2020). The low nutrient concentration in the rivers object of our study and the lower light intensity might explain the higher decomposition observed in the downstream reaches. Stable flows are also related to increased bank stability and encroachment of riparian vegetation which is related to increased canopy cover (Hadley et al., 2018). An increased canopy is often related to increase in shaded areas, where the ratio between heterotrophic and autotrophic activity can lean in favor of the former, enhancing ecosystem respiration over primary production (Burrell et al., 2014).

If flow stability can positively influence the organic matter decomposition, on the other hand, biofilm communities might not be affected in the same way. There was no difference in the microbial community composition between upstream and downstream sites, whereas the communities seemed to be more strictly related to the source catchment, which was the driver explaining most of the variation among them, however, high natural variability might have also masked the hydropower impact. In their study, Wagner et al. (2015) reported that differences in light intensity did don't show significant changes for microbial diversity, even though changes in the functional processes were occurring, mirroring the results shown in our study.

Ceramic tiles were selected as a standard substrate to reduce the variability due to the riverbed geology, however this choice may have confounded the results. The preference for the texture of

the tiles' substrate might have acted as a filter for certain taxa, homogenizing the communities and preventing the capture of the true natural variability of the prokaryotic communities in the upstream and downstream river reaches.

Some limitations of this study were related to it being only a snapshot in time, so it was not possible to capture the seasonal variability which could characterize upstream and downstream communities throughout the year. Replication within each site was also difficult due to the slow growth of biofilms during the period of exposure, thus proper number of replicates would also help in determining the patterns of microbial community structure between impacted and unimpacted sites.

Response to different flow regime scenarios

The evaluation of different flow scenarios on the prokaryotic community structure and functioning was the central topic of **Paper III** which was carried out in a mesocosm experiment consisting of sixteen flumes at the Birmingham University. The results of this study reveal how flow alterations can lead to significant differences and dynamics in the prokaryotic community composition and also the processes carried out by these communities, leading to unpredictable impacts on the relationships within the biofilm components and the higher trophic levels (Juvigny-Khenafou et al., 2021).

In this experimental study we found similar prokaryotic taxa to be dominant in the biofilm communities, with a prevalence of Proteobacteria, followed by Planctomycetota and Bacteroidota, Verrucomicrobiota and Actinobacteriota, similar results to what we found for **paper I** and **paper II** and to findings from other studies (Battin et al., 2016; Juvigny-Khenafou et al., 2021; Romaní et al., 2017). Among the Proteobacteria, we found the Alphaproteobacteria to be the dominant class, followed by the Gammaproteobacteria and Planctomycetes, Verrucomicrobiae, Actinobacteria and Bacteroidia, mirroring the results from the prokaryotic analysis from **paper I** and **II.**

Species turnover or beta diversity was the parameter most influenced by the different flow regimes implemented in the flumes. Contrary, we did not detect any significant change in the alpha diversity measures, apart from a significant difference highlighted between the Pielu's evenness values between natural and homogenized treatments from the third sampling occasion. The natural treatments showed the highest evenness values among all treatments, while the homogenized had the lowest. This difference was significant for the riffle habitats. Possibly, in the long term, the species succession would shift towards species more tolerant to stable flows in the homogenized treatments, while the flow variability of natural flow regime would maintain the species diversity without favoring any specific taxa. Other studies (Sabater et al., 2018; González & Elosegi, 2021) showed how human activity can drive water stress reducing the diversity and evenness of biological communities in river reaches impacted by flow stabilization and reduction in natural flow variability. The treatment showing the most striking discrepancies when compared to the natural flow treatment, which was our control, was the drought treatment, consisting in a flow reduction of 60%. Both the community composition and the organic matter degradation were consistently affected by the reduction in the water discharge throughout the exposure period. A study by Juvigny-Khenafou et al. (2021), also working with mesocosms experiments and microbial communities, reported similar results. In their study, the microbial community composition was consistently affected when treated with flow reduction and nutrient enrichments. The main drivers for the community dynamics were flow reduction and sediment addition (Juvigny-Khenafou et al. 2021). Moreover, they also showed that functional genes were affected in the same manner as community composition, revealing a coupled behavior among structure and function of the microbial community similar to the coupling we found in the drought treatments. In our study the decomposition of organic matter was significantly higher in the drought treatments, for both, fine and coarse mesh bags, highlighting an enhanced heterotrophic activity within the flumes with reduced water discharge. An increase in heterotrophic activity due to droughts has been reported in previous studies (Aristi et al., 2014; Ponsatí et al., 2015). Aristi et

al. (2014) showed how flow regulation due to dams can have severe effects on the functional processes of river ecosystems. The accumulation of carbon sources, both allochthonous and autochthonous, were enhanced by the reduced discharge and flow variability below dams, which represents a similar scenario to our drought treatment. They showed how both, gross primary production and ecosystem respiration increased respectively by 59% and 75% due to the low flow and reduced scouring on the riverbed. Higher metabolic rates due to high water retention time and faster processing of organic matter in the downstream reaches might have significant repercussions on the carbon cycling, reducing the flux of organic carbon towards the downstream reaches, increasing its transformation to inorganic form. Faster organic matter processing may increase the self-purification capacity of the rivers experiencing droughts (Acuña et al., 2013) but it may also increase the carbon dioxide emissions from river systems, reducing the ability of rivers to act as carbon sinks (Battin et al., 2009). Moreover, the fast organic matter processing within river systems may alter the flux of energy to the coastal zones, resulting in a depletion of bioavailable organic matter for the marine communities (Bernal & Sabater, 2012; von Schiller et al., 2017). This is particularly relevant, as prediction of increases in temperatures by 1-3 °C by mid-21st century and between 2-5 °C by late-21st century (Field et al., 2012), due to climate change, may consistently affect river ecosystems, extending the droughts periods and increasing water temperatures (Timoner et al., 2014) leading to significant impacts on the structure and functioning of river ecosystems.

The results from both **paper II** and **III** revealed how ecosystem processes, such as decomposition and cycling of carbon, might be affected by flow alterations due to hydropower impact in river ecosystems and more generally to modifications of the natural flow regime. Different results were observed in terms of the prokaryotic community structure, which in natural environments (**paper II**) did not show significant differences comparing upstream and downstream sites. However, in the mesocosms experiments the effects on the species turnover were clear. Controlling for the confounding variables in the mesocosms experiment, allowed for a reduction in the variability

naturally occurring in river ecosystems, which in **paper II,** might have masked the differences due to the hydropower impact. Although, different responses from structure and function of microbial communities to environmental stressors has been observed before and ascribed to adaptive responses (Fellows et al., 2006).

From the experiments in the mesocosms implementing different flow regimes, we can clearly see that alterations to the natural flow variability can have significant effects on both, community dynamic and ecological processes, corroborating the findings on the organic matter breakdown from **paper II.**

From the results of the drought treatments in **paper III**, simulating future climate change scenarios with enhanced water shortage, we can see how modifications to the natural flow regime could result in significant changes to the ecosystem processes and river biodiversity, leading to unpredictable impacts on the energy flux and biogeochemical processes from the river to the coastal areas as already suggested by Palmer & Ruhi in 2019.

Developing prokaryotic bioindicators for classification of the ecological status of river ecosystems (Paper I-IV)

In **paper I** we proposed a new methodology to identify prokaryotes at different taxonomic levels which may be used as bioindicators for specific environmental variables. The ASVs having the highest variation in the relative abundances and occurring more often in the sampling sites were identified as the most informative as they would react to specific environmental drivers in most pronounced way. The taxonomic classification was performed at at different taxonomic resolutions including phylum, class and genus levels. Redundancy Analyses were performed by using the Bray-Curtis dissimilarity matrices for each of the taxonomic levels across the main environmental gradients. Specific relationships were detected among specific taxa such as Gammaproteobacteria and high loads of Total Carbon and Nitrogen and the class of Alphaproteobacteria and low pH levels. While at genus level, *Pseudomonas* was the most relevant

taxa and it was strictly linked to high values of total Nitrogen and Carbon, for which also *Acinetobacter* showed strong relationships. *Sphingomonas,* genus belonging to the class of Alphaproteobacteria, seemed to be heavily influenced by the pH gradient and nutrient concentrations, being significantly more abundant in the sampling sites characterized by higher acidity and overall low levels of nutrients.

This approach was compared to machine learning methodologies in **paper IV** where, by using existing water quality classification based on the Saprobic Index from macroinvertebrate samples and chlorophyll-a concentration as a proxy for eutrophication, we detect informative ASVs for water quality predictions. Results from the prevalence/variance analysis and the spatio-temporal approach (based on machine learning) pointed to ASVs belonging to the same overarching groups as good biological indicators, highlighting the relevance of specific prokaryotic classes in terms of water quality assessment for river systems. The drifting behavior of the prokaryotes in the river ecosystem allowed for prediction of the water quality at specific sites by looking at the ASVs abundances in the upstream sites. The ability to predict water quality without having data for a specific site could allow to analyze river reaches where the collection of data is not possible. Moreover, the knowledge related to the number of upstream sampling sites needed for an accurate prediction of the water quality for a specific site will be helpful for water policy makers to implement restoration projects and activities on a wider area. This should cover not only the stretch of interest but also upstream sites where the source of the decrease in water quality is located.

To conclude, the inclusion of prokaryotes in the biomonitoring network could complement the already existing methods and, in some cases, the use of molecular techniques joint with machine learning algorithms could improve the harmonization and intercalibration of the ecological status assessment in different countries, as it can be standardized and it would not suffer from discrepancies related to morphological identification and different systems for the evaluation of the water quality status.

Perspectives

The world is shifting every day towards more sustainable sources of energy and hydropower is regarded as one of the main solutions to reduce the greenhouse gasses emission around the globe. However, hydroelectric energy production does not come without environmental costs and t many rivers worldwide are affected by hydromorphological alterations due to human exploitation for different purposes. The result of this overexploitation of river ecosystems is an increasing loss of biodiversity and a general decrease of the ecological integrity of these environments. In this thesis, I analyzed the effects of human induced flow alterations on a still poorly understood biological component, namely the prokaryotic assemblages living in riverine biofilms.

The results from the four papers object of this thesis suggest an additional detrimental effect of flow regime modifications to the smallest organisms inhabiting rivers. Despite their dimension, prokaryotes are crucial player for the ecosystems, being at the base of the food webs, providing food for higher trophic levels and playing a central role in several biogeochemical processes. Their sensitivity to ecological shifts makes them a perfect fit among the bioindicators currently in use and, the exploitation of new molecular methodologies such as metabarcoding, would make them easily comparable across countries reducing the risks of misinterpretation due to the morphological identification. International frameworks for the environmental impact assessment such as the Water Framework Directive are still neglecting these microbial communities for the classification of the ecological status of natural ecosystems. However, we stress the need for inclusion of microbes in the routine of the current monitoring networks as they can act as early warnings for pollution, eutrophication and even hydro-morphological alteration as seen from this thesis results. Moreover, an increased use of ecosystem functioning proxies should be implemented since biodiversity measures alone cannot provide a comprehensive overview of the ecological status of a water body. Ecosystem processes such as organic matter decomposition can

be a good indicator of changes in the ecological status of rivers, especially when looking at climate change scenarios. Although, more research embracing community structure and functioning at the same time, is still needed to broaden our understanding of the relationships between flow alterations and river ecology.

Overall, this thesis highlights the risks linked to flow modifications in relation to the structure and functioning of microbial communities, which have been overlooked in studies relating to flow management and river ecology so far. The importance of these organisms for the ecology of river ecosystems should be an incentive for their inclusion in the environmental monitoring network worldwide.

References

Acuña, V., Díez, J. R., Flores, L., Meleason, M., & Elosegi, A. (2013). Does it make economic sense to restore rivers for their ecosystem services? *Journal of Applied Ecology*, *50*(4), 988–997. https://doi.org/10.1111/1365-2664.12107

Alfredsen, K., Harby, A., Linnansaari, T., & Ugedal, O. (2012). Development of an Inflow-Controlled Environmental Flow Regime for a Norwegian River. *River Research and Applications*, *28*(6), 731–739. https://doi.org/10.1002/rra.1550

Allan, J. D., Castillo, M. M., & Capps, K. A. (2021). *Stream Ecology: Structure and Function of Running Waters*. Springer International Publishing. https://doi.org/10.1007/978-3-030-61286-3

Aristi, I., Arroita, M., Larrañaga, A., Ponsatí, L., Sabater, S., Schiller, D. von, Elosegi, A., & Acuña, V. (2014). Flow regulation by dams affects ecosystem metabolism in Mediterranean rivers. *Freshwater Biology*, *59*(9), 1816–1829. https://doi.org/10.1111/fwb.12385

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

Azam, F., & Malfatti, F. (2007). Microbial structuring of marine ecosystems. *Nature Reviews Microbiology*, *5*(10), 782–791. https://doi.org/10.1038/nrmicro1747

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

Bakken, T. H., Sundt, H., Ruud, A., & Harby, A. (2012). Development of Small Versus Large Hydropower in Norway– Comparison of Environmental Impacts. *Energy Procedia*, *20*, 185–199. https://doi.org/10.1016/j.egypro.2012.03.019

Batalla, R. J., Gómez, C. M., & Kondolf, G. M. (2004). Reservoir-induced hydrological changes in the Ebro River basin (NE Spain). *Journal of Hydrology*, *290*(1–2), 117–136. https://doi.org/10.1016/j.jhydrol.2003.12.002

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., Luyssaert, S., Kaplan, L. A., Aufdenkampe, A. K., Richter, A., & Tranvik, L. J. (2009). The boundless carbon cycle. *Nature Geoscience*, *2*(9), 598–600. https://doi.org/10.1038/ngeo618

Bernal, S., & Sabater, F. (2012). Changes in discharge and solute dynamics between hillslope and valley-bottom intermittent streams. *Hydrology and Earth System Sciences*, *16*(6), 1595–1605. https://doi.org/10.5194/hess-16-1595-2012

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., Moeseneder, M. M., Arrieta, J. M., Herndl, G. J., & Peduzzi, P. (2005). Complexity of Bacterial Communities in a River-Floodplain System (Danube, Austria). *Applied and Environmental Microbiology*, *71*(2), 609–620. https://doi.org/10.1128/AEM.71.2.609-620.2005

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

Bloem, J., & Breure, A. M. (2003). Chapter 8 Microbial indicators. In *Trace Metals and other Contaminants in the Environment* (Vol. 6, pp. 259–282). Elsevier. https://doi.org/10.1016/S0927-5215(03)80138-8

Böhme, A., Risse-Buhl, U., & Küsel, K. (2009). Protists with different feeding modes change biofilm morphology. *FEMS Microbiology Ecology*, *69*(2), 158–169. https://doi.org/10.1111/j.1574-6941.2009.00710.x

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*. https://doi.org/10.3389/fmicb.2011.00275

Burgos-Caraballo, S., Cantrell, S. A., & Ramírez, A. (2014). Diversity of benthic biofilms along a land use gradient in tropical headwater streams, Puerto Rico. *Microbial Ecology*, *68*(1), 47–59. https://doi.org/10.1007/s00248-014-0401-x

Burrell, T. K., O'Brien, J. M., Graham, S. E., Simon, K. S., Harding, J. S., & McIntosh, A. R. (2014). Riparian shading mitigates stream eutrophication in agricultural catchments. *Freshwater Science*, *33*(1), 73–84. https://doi.org/10.1086/674180

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

Cordier, T., Alonso-Sáez, L., Apothéloz-Perret-Gentil, L., Aylagas, E., Bohan, D. A., Bouchez, A., Chariton, A., Creer, S., Frühe, L., Keck, F., Keeley, N., Laroche, O., Leese, F., Pochon, X., Stoeck, T., Pawlowski, J., & Lanzén, A. (2020). Ecosystems monitoring powered by environmental genomics: A review of current strategies with an implementation roadmap. *Molecular Ecology*, *n/a*(n/a). https://doi.org/10.1111/mec.15472

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

Dolores Bejarano, M., & Sordo-Ward, Á. (2011). Riparian woodland encroachment following flow regulation: A comparative study of Mediterranean and Boreal streams. *Knowledge and Management of Aquatic Ecosystems*, *402*, 20. https://doi.org/10.1051/kmae/2011059

Dopheide, A., Lear, G., Stott, R., & Lewis, G. (2008). Molecular characterization of ciliate diversity in stream biofilms. *Applied and Environmental Microbiology*, *74*(6), 1740–1747. https://doi.org/10.1128/aem.01438-07

Dorber, M., Mattson, K. R., Sandlund, O. T., May, R., & Verones, F. (2019). Quantifying net water consumption of Norwegian hydropower reservoirs and related aquatic biodiversity impacts in Life Cycle Assessment. *Environmental Impact Assessment Review*, *76*, 36–46. https://doi.org/10.1016/j.eiar.2018.12.002

Dudgeon, D. (2019). Multiple threats imperil freshwater biodiversity in the Anthropocene. *Current Biology*, *29*(19), R960–R967. https://doi.org/10.1016/j.cub.2019.08.002

Dudgeon, D., Arthington, A. H., Gessner, M. O., Kawabata, Z.-I., Knowler, D. J., Lévêque, C., Naiman, R. J., Prieur-Richard, A.-H., Soto, D., Stiassny, M. L. J., & Sullivan, C. A. (2006). Freshwater biodiversity: Importance, threats, status and conservation challenges. *Biological Reviews*, *81*(2), 163–182. https://doi.org/10.1017/S1464793105006950

Eiler, A., Hayakawa, D. H., Church, M. J., Karl, D. M., & Rappé, M. S. (2009). Dynamics of the SAR11 bacterioplankton lineage in relation to environmental conditions in the oligotrophic North Pacific subtropical gyre. *Environmental Microbiology*, *11*(9), 2291–2300. https://doi.org/10.1111/j.1462-2920.2009.01954.x

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

Field, C. B., Barros, V., Stocker, T. F., & Dahe, Q. (Eds.). (2012). *Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation: Special Report of the*

Intergovernmental Panel on Climate Change. Cambridge University Press. https://doi.org/10.1017/CBO9781139177245

Findlay, S., Tank, J., Dye, S., Valett, H. M., Mulholland, P. J., McDowell, W. H., Johnson, S. L., Hamilton, S. K., Edmonds, J., Dodds, W. K., & Bowden, W. B. (2002). A Cross-System Comparison of Bacterial and Fungal Biomass in Detritus Pools of Headwater Streams. *Microbial Ecology*, *43*(1), 55–66. https://doi.org/10.1007/s00248-001-1020-x

Fortunato, C. S., Eiler, A., Herfort, L., Needoba, J. A., Peterson, T. D., & Crump, B. C. (2013). Determining indicator taxa across spatial and seasonal gradients in the Columbia River coastal margin. *The ISME Journal*, *7*(10), 1899–1911. https://doi.org/10.1038/ismej.2013.79

Foucreau, N., Puijalon, S., Hervant, F., & Piscart, C. (2013). Effect of leaf litter characteristics on leaf conditioning and on consumption by Gammarus pulex. *Freshwater Biology*, *58*(8), 1672–1681. https://doi.org/10.1111/fwb.12158

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

Gessner, M. O., & Chauvet, E. (2002). A CASE FOR USING LITTER BREAKDOWN TO ASSESS FUNCTIONAL STREAM INTEGRITY. *Ecological Applications*, *12*(2), 498–510. https://doi.org/10.1890/1051-0761(2002)012[0498:ACFULB]2.0.CO;2

Gessner, M. O., Chauvet, E., & Dobson, M. (1999). A Perspective on Leaf Litter Breakdown in Streams. *Oikos*, *85*(2), 377–384. https://doi.org/10.2307/3546505

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. *Environmental Microbiology Reports*, *9*(2), 129–143. https://doi.org/10.1111/1758-2229.12509

González, J. M., & Elosegi, A. (2021). Water abstraction reduces taxonomic and functional diversity of stream invertebrate assemblages. *Freshwater Science*, *40*(3), 524–536. https://doi.org/10.1086/716201

Hadley, D. R., Grams, P. E., & Kaplinski, M. A. (2018). Quantifying geomorphic and vegetation change at sandbar campsites in response to flow regulation and controlled floods, Grand Canyon National Park, Arizona. *River Research and Applications*, *34*(9), 1208–1218. https://doi.org/10.1002/rra.3349

Hein, T., Baranyi, C., Herndl, G. J., Wanek, W., & Schiemer, F. (2003). Allochthonous and autochthonous particulate organic matter in floodplains of the River Danube: The importance of hydrological connectivity. *Freshwater Biology*, *48*(2), 220–232. https://doi.org/10.1046/j.1365-2427.2003.00981.x

Howard-Parker, B., White, B., Halvorson, H. M., & Evans-White, M. A. (2020). Light and dissolved nutrients mediate recalcitrant organic matter decomposition via microbial priming in experimental streams. *Freshwater Biology*, *65*(7), 1189–1199. https://doi.org/10.1111/fwb.13503

Jezbera, J., Jezberová, J., Koll, U., Horňák, K., Šimek, K., & Hahn, M. W. (2012). Contrasting trends in distribution of four major planktonic betaproteobacterial groups along a pH gradient of epilimnion of 72 freshwater habitats. *FEMS Microbiology Ecology*, *81*(2), 467–479. https://doi.org/10.1111/j.1574-6941.2012.01372.x

Johnsen, A. R., Hausner, M., Schnell, A., & Wuertz, S. (2000). Evaluation of Fluorescently Labeled Lectins for Noninvasive Localization of Extracellular Polymeric Substances inSphingomonas Biofilms. *Applied and Environmental Microbiology*, *66*(8), 3487–3491. https://doi.org/10.1128/AEM.66.8.3487-3491.2000

Jungwirth, M., Muhar, S., & Schmutz, S. (Eds.). (2000). *Assessing the Ecological Integrity of Running Waters: Proceedings of the International Conference, held in Vienna, Austria, 9–11 November 1998*. Springer Netherlands. https://doi.org/10.1007/978-94-011-4164-2

Juvigny-Khenafou, N. P. D., Piggott, J. J., Atkinson, D., Zhang, Y., Wu, N., & Matthaei, C. D. (2021). Fine sediment and flow velocity impact bacterial community and functional profile more than nutrient enrichment. *Ecological Applications*, *31*(1), e02212. https://doi.org/10.1002/eap.2212

Kaspari, M. (2004). Using the Metabolic Theory of Ecology to Predict Global Patterns of Abundance. *Ecology*, *85*(7), 1800–1802. https://doi.org/10.1890/03-0682

Kirchman, D. L. (2002). The ecology of Cytophaga-Flavobacteria in aquatic environments. *FEMS Microbiology Ecology*, *39*(2), 91–100. https://doi.org/10.1111/j.1574-6941.2002.tb00910.x

Lawrence, J. R., Scharf, B., Packroff, G., & Neu, T. R. (2002). Microscale Evaluation of the Effects of Grazing by Invertebrates with Contrasting Feeding Modes on River Biofilm Architecture and Composition. *Microbial Ecology*, *44*(3), 199–207. https://doi.org/10.1007/s00248-001-1064-y

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

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

Mateo, P., Leganés, F., Perona, E., Loza, V., & Fernández-Piñas, F. (2015). Cyanobacteria as bioindicators and bioreporters of environmental analysis in aquatic ecosystems.

Biodiversity and Conservation, *24*(4), 909–948. https://doi.org/10.1007/s10531-015-0903-y

Meißner, T., Schütt, M., Sures, B., & Feld, C. K. (2018). Riverine regime shifts through reservoir dams reveal options for ecological management. *Ecological Applications*, *28*(7), 1897–1908. https://doi.org/10.1002/eap.1786

Moe, T. F., Brysting, A. K., Andersen, T., Schneider, S. C., Kaste, Ø., & Hessen, D. O. (2013). Nuisance growth of Juncus bulbosus: The roles of genetics and environmental drivers tested in a large-scale survey. *Freshwater Biology*, *58*(1), 114–127. https://doi.org/10.1111/fwb.12043

Newton, R. J., Jones, S. E., Eiler, A., McMahon, K. D., & Bertilsson, S. (2011). A Guide to the Natural History of Freshwater Lake Bacteria. *Microbiology and Molecular Biology Reviews*, *75*(1), 14–49. https://doi.org/10.1128/MMBR.00028-10

Nilsson, C., Reidy, C. A., Dynesius, M., & Revenga, C. (2005). Fragmentation and Flow Regulation of the World's Large River Systems. *Science*, *308*(5720), 405–408. JSTOR.

Nunes, M., & Adams, J. B. (2014). Responses of primary producers to mouth closure in the temporarily open/closed Great Brak Estuary in the warm-temperate region of South Africa. *African Journal of Aquatic Science*, *39*(4), 387–394. https://doi.org/10.2989/16085914.2014.980773

Palmer, M., & Ruhi, A. (2019). Linkages between flow regime, biota, and ecosystem processes: Implications for river restoration. *Science*, *365*(6459), eaaw2087. https://doi.org/10.1126/science.aaw2087

Pawlowski, J., Kelly-Quinn, M., Altermatt, F., Apothéloz-Perret-Gentil, L., Beja, P., Boggero, A., Borja, A., Bouchez, A., Cordier, T., Domaizon, I., Feio, M. J., Filipe, A. F., Fornaroli, R., Graf, W., Herder, J., van der Hoorn, B., Iwan Jones, J., Sagova-Mareckova, M., Moritz, C., … 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

Poff, N. L., Allan, J. D., Bain, M. B., Karr, J. R., Prestegaard, K. L., Richter, B. D., Sparks, R. E., & Stromberg, J. (1997). The natural flow regime: A paradigm for river conservation and restoration. *BioScience*, *47*(11), 769–784. https://doi.org/10.2307/1313099

Ponsatí, L., Acuña, V., Aristi, I., Arroita, M., García-Berthou, E., Schiller, D. von, Elosegi, A., & Sabater, S. (2015). Biofilm Responses to Flow Regulation by Dams in Mediterranean Rivers. *River Research and Applications*, *31*(8), 1003–1016. https://doi.org/10.1002/rra.2807

Prats, J., Val, R., Armengol, J., & Dolz, J. (2010). Temporal variability in the thermal regime of the lower Ebro River (Spain) and alteration due to anthropogenic factors. *Journal of Hydrology*, *387*(1), 105–118. https://doi.org/10.1016/j.jhydrol.2010.04.002

Reid, A. J., Carlson, A. K., Creed, I. F., Eliason, E. J., Gell, P. A., Johnson, P. T. J., Kidd, K. A., MacCormack, T. J., Olden, J. D., Ormerod, S. J., Smol, J. P., Taylor, W. W., Tockner, K., Vermaire, J. C., Dudgeon, D., & Cooke, S. J. (2019). Emerging threats and persistent conservation challenges for freshwater biodiversity. *Biological Reviews*, *94*(3), 849–873. https://doi.org/10.1111/brv.12480

Richter, B. D., Postel, S., Revenga, C., Scudder, T., Lehner, B., Churchill, A., & Chow, M. (2010). *Lost in Development's Shadow: The Downstream Human Consequences of Dams*. *3*(2), 29.

Risse-Buhl, U., Trefzger, N., Seifert, A.-G., Schönborn, W., Gleixner, G., & Küsel, K. (2012). Tracking the autochthonous carbon transfer in stream biofilm food webs. *FEMS Microbiology Ecology*, *79*(1), 118–131. https://doi.org/10.1111/j.1574-6941.2011.01202.x

Rivaes, R., Rodríguez-González, P. M., Albuquerque, A., Pinheiro, A. N., Egger, G., & Ferreira, M. T. (2015). Reducing river regulation effects on riparian vegetation using flushing flow

regimes. *Ecological Engineering*, *81*, 428–438. https://doi.org/10.1016/j.ecoleng.2015.04.059

Romaní, A. M., Amalfitano, S., Artigas, J., Fazi, S., Sabater, S., Timoner, X., Ylla, I., & Zoppini, A. (2013). Microbial biofilm structure and organic matter use in mediterranean streams. *Hydrobiologia*, *719*(1), 43–58. https://doi.org/10.1007/s10750-012-1302-y

Romaní, A. M., Chauvet, E., Febria, C., Mora-Gómez, J., Risse-Buhl, U., Timoner, X., Weitere, M., & Zeglin, L. (2017). The Biota of Intermittent Rivers and Ephemeral Streams: Prokaryotes, Fungi, and Protozoans. In *Intermittent Rivers and Ephemeral Streams* (pp. 161–188). Elsevier. https://doi.org/10.1016/B978-0-12-803835-2.00009-7

Sabater, S., Timoner, X., Borrego, C., & Acuña, V. (2016). Stream Biofilm Responses to Flow Intermittency: From Cells to Ecosystems. *Frontiers in Environmental Science*, *4*. https://doi.org/10.3389/fenvs.2016.00014

Sabater, S., Bregoli, F., Acuña, V., Barceló, D., Elosegi, A., Ginebreda, A., Marcé, R., Muñoz, I., Sabater-Liesa, L., & Ferreira, V. (2018). Effects of human-driven water stress on river ecosystems: A meta-analysis. *Scientific Reports*, *8*(1), 11462. https://doi.org/10.1038/s41598-018-29807-7

Salcher, M. M., Ewert, C., Šimek, K., Kasalický, V., & Posch, T. (2016). Interspecific competition and protistan grazing affect the coexistence of freshwater betaproteobacterial strains. *FEMS Microbiology Ecology*, *92*(2). https://doi.org/10.1093/femsec/fiv156

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

Schneider, S. C., Fosholt Moe, T., Hessen, D. O., & Kaste, Ø. (2013). Juncus bulbosus nuisance growth in oligotrophic freshwater ecosystems: Different triggers for the same phenomenon

in rivers and lakes? *Aquatic Botany*, *104*, 15–24. https://doi.org/10.1016/j.aquabot.2012.10.001

Schneider, S. C., Sample, J. E., Moe, J. S., Petrin, Z., Meissner, T., & Hering, D. (2018). Unravelling the effect of flow regime on macroinvertebrates and benthic algae in regulated versus unregulated streams. *Ecohydrology*, *11*(7), e1996. https://doi.org/10.1002/eco.1996

Smolar-Žvanut, N., & Mikoš, M. (2014). The impact of flow regulation by hydropower dams on the periphyton community in the Soča River, Slovenia. *Hydrological Sciences Journal*, *59*(5), 1032–1045. https://doi.org/10.1080/02626667.2013.834339

Thorstad, E. B., Okland, F., Kroglund, F., & Jepsen, N. (2003). Upstream migration of Atlantic salmon at a power station on the River Nidelva, Southern Norway. *Fisheries Management and Ecology*, *10*(3), 139–146. https://doi.org/10.1046/j.1365-2400.2003.00335.x

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

Timoner, X., Borrego, C. M., Acuña, V., & Sabater, S. (2014). The dynamics of biofilm bacterial communities is driven by flow wax and wane in a temporary stream. *Limnology and Oceanography*, *59*(6), 2057–2067. https://doi.org/10.4319/lo.2014.59.6.2057

Vannote, R. L., Minshall, G. W., Cummins, K. W., Sedell, J. R., & Cushing, C. E. (1980). The River Continuum Concept. *Canadian Journal of Fisheries and Aquatic Sciences*, *37*(1), 130–137. https://doi.org/10.1139/f80-017

Visser, P. M., Verspagen, J. M. H., Sandrini, G., Stal, L. J., Matthijs, H. C. P., Davis, T. W., Paerl, H. W., & Huisman, J. (2016). How rising CO2 and global warming may stimulate harmful cyanobacterial blooms. *Harmful Algae*, *54*, 145–159. https://doi.org/10.1016/j.hal.2015.12.006

von Schiller, D., Bernal, S., Dahm, C. N., & Martí, E. (2017). Nutrient and Organic Matter Dynamics in Intermittent Rivers and Ephemeral Streams. In *Intermittent Rivers and Ephemeral Streams* (pp. 135–160). Elsevier. https://doi.org/10.1016/B978-0-12-803835-2.00006-1

Wagner, K., Besemer, K., Burns, N. R., Battin, T. J., & Bengtsson, M. M. (2015). Light availability affects stream biofilm bacterial community composition and function, but not diversity. *Environmental Microbiology*, *17*(12), 5036–5047. https://doi.org/10.1111/1462-2920.12913

Wang, Y., Zhang, N., Wang, D., Wu, J., & Zhang, X. (2018). Investigating the impacts of cascade hydropower development on the natural flow regime in the Yangtze River, China. *The Science of the Total Environment*, *624*, 1187–1194. https://doi.org/10.1016/j.scitotenv.2017.12.212

Wey, J. K., Jürgens, K., & Weitere, M. (2012). Seasonal and Successional Influences on Bacterial Community Composition Exceed That of Protozoan Grazing in River Biofilms. *Applied and Environmental Microbiology*, *78*(6), 2013–2024. https://doi.org/10.1128/AEM.06517-11

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

Wohl, E. (2018). *Sustaining River Ecosystems and Water Resources*. Springer International Publishing. https://doi.org/10.1007/978-3-319-65124-8

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

Wright, R. F., Couture, R.-M., Christiansen, A. B., Guerrero, J.-L., Kaste, Ø., & Barlaup, B. T. (2017). Effects of multiple stresses hydropower, acid deposition and climate change on water chemistry and salmon populations in the River Otra, Norway. *Science of The Total Environment*, *574*, 128–138. https://doi.org/10.1016/j.scitotenv.2016.09.044

Zarfl, C., Berlekamp, J., He, F., Jähnig, S. C., Darwall, W., & Tockner, K. (2019). Future large hydropower dams impact global freshwater megafauna. *Scientific Reports*, *9*(1), 18531. https://doi.org/10.1038/s41598-019-54980-8

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

Websites

Energifakta Norway. 2021. https://energifaktanorge.no/ last visited 25.06.2021

NVE. 2021. https://www.nve.no/konsesjonssaker/konsesjonsbehandling-av-vannkraft/last visited 25.06.2021

SSB. 2021. https://www.ssb.no/energi-og-industri/artikler-og-publikasjoner/pa-verdenstoppen-i-bruk-av-strom last visited 25.06.2021

Water Resources Act – Regjering.no https://www.regjeringen.no/globalassets/upload/oed/vedlegg/lover-og-reglement/act_no_82_of_24_november_2000.pdf last visited 28.06.2021

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

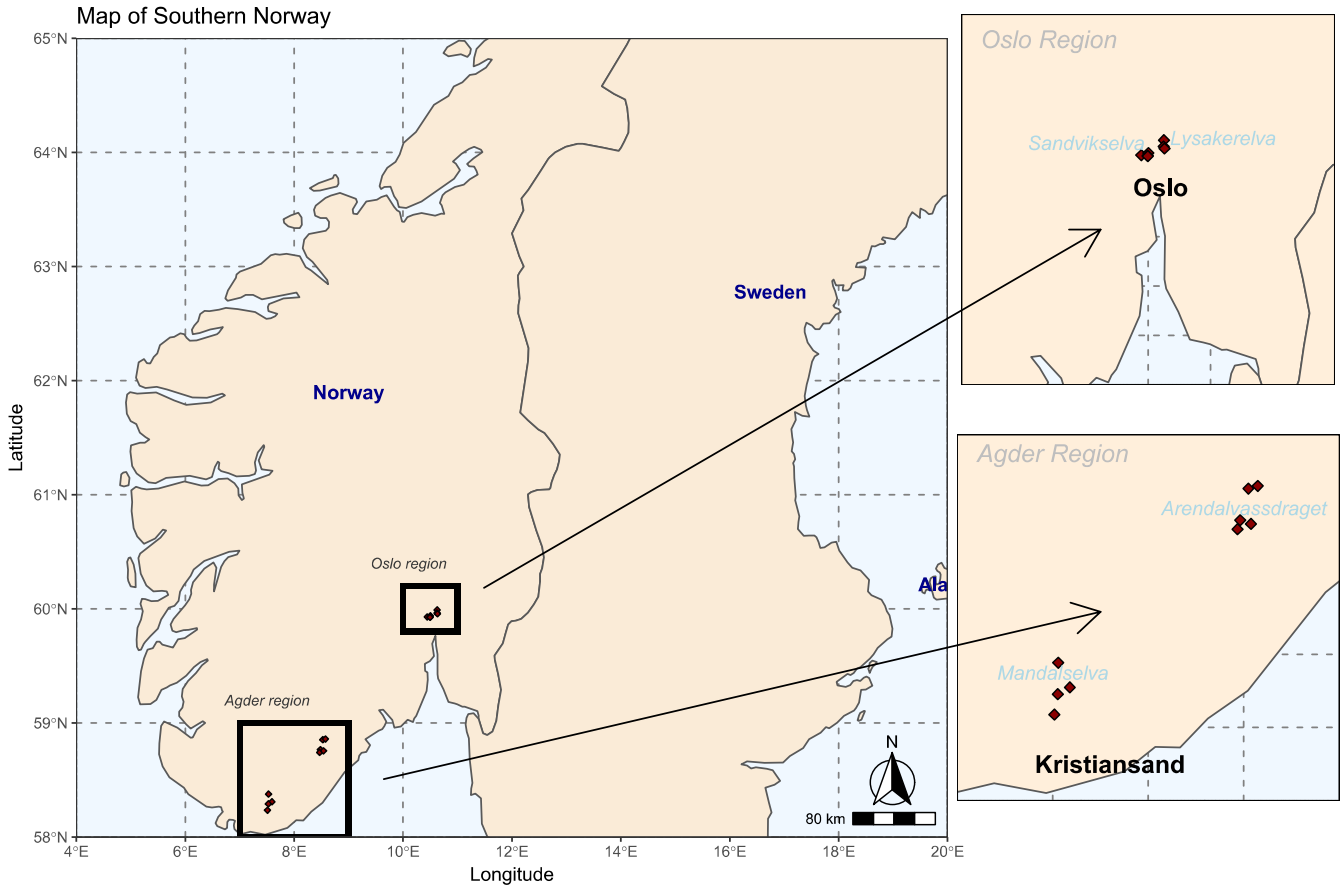


FIGURE 1 Map of Norway and location of the sampling sites in the two regions. Highlighted in the small pictures are the four catchments

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

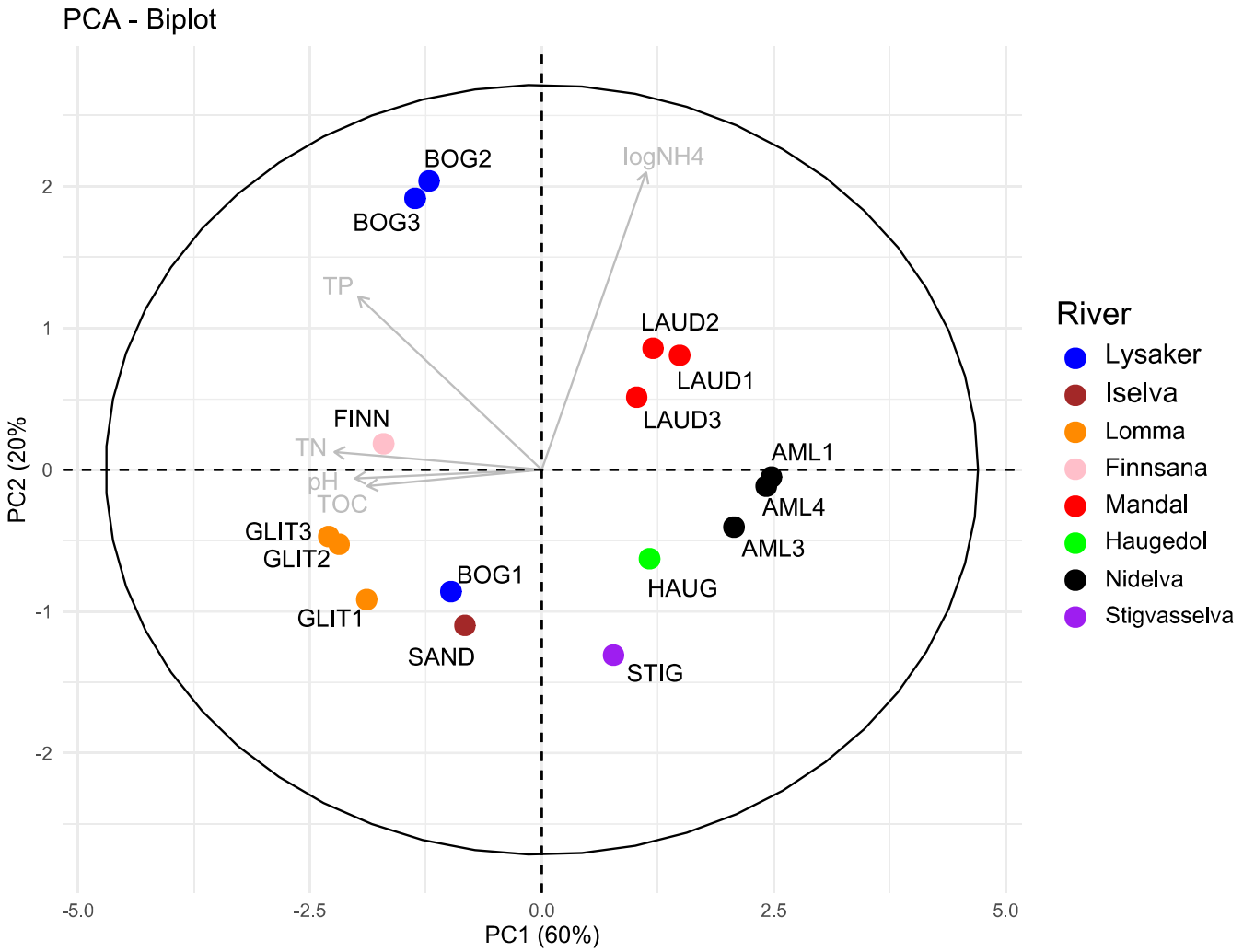


FIGURE 2 PCA with the relevant standardized environmental variables. Length of vectors is proportional to the contribution of the variable to the principal components. Colours correspond to the sampled rivers

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

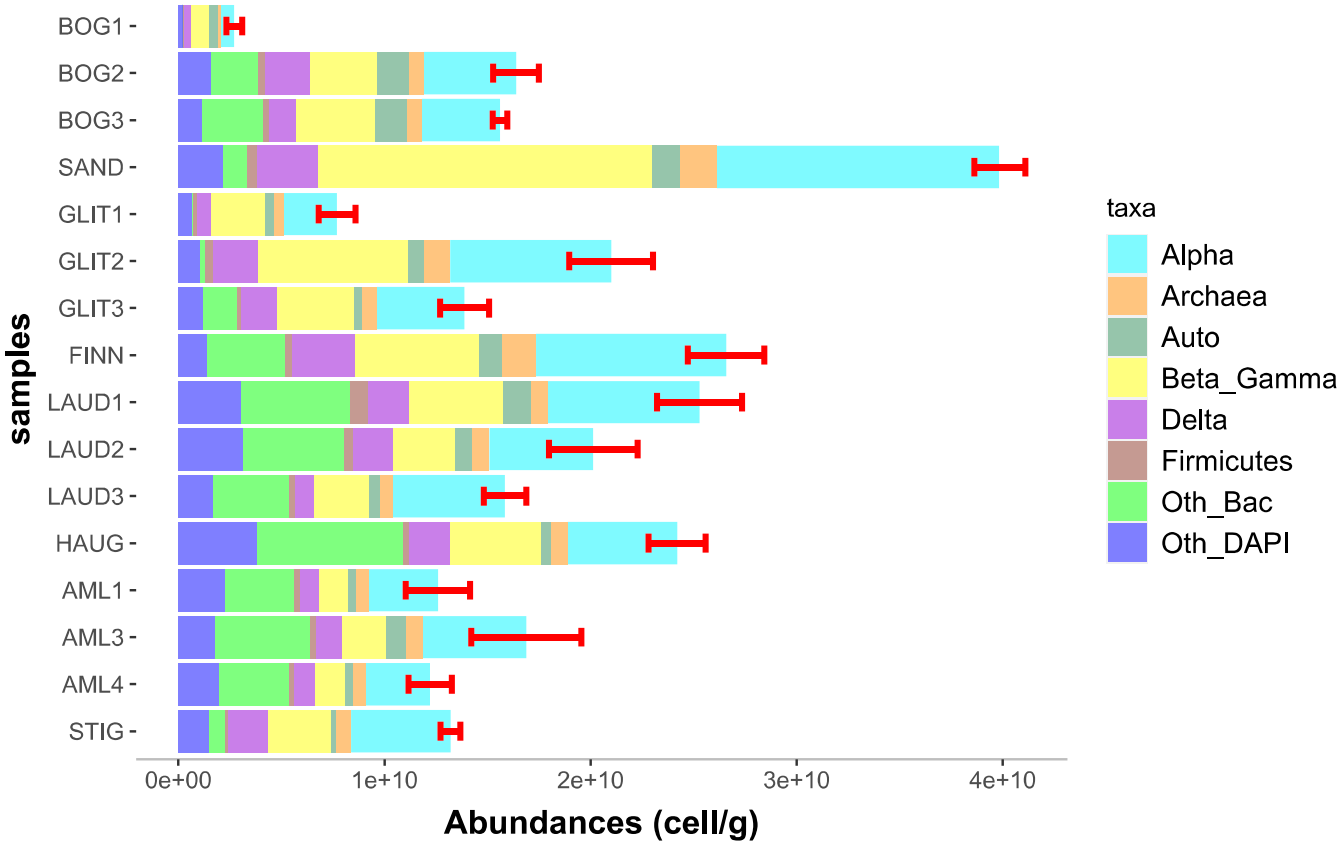


FIGURE 3 Absolute cell numbers for the prokaryotic taxa analysed by CARD-FISH. The key shows the classes Proteobacteria (Alpha-, Beta- and Gamma-[joint abundances] and Delta-); the phylum Firmicutes (Firmicutes); the autofluorescent cells (Auto), which correspond to photosynthetic prokaryotes; archaeal 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

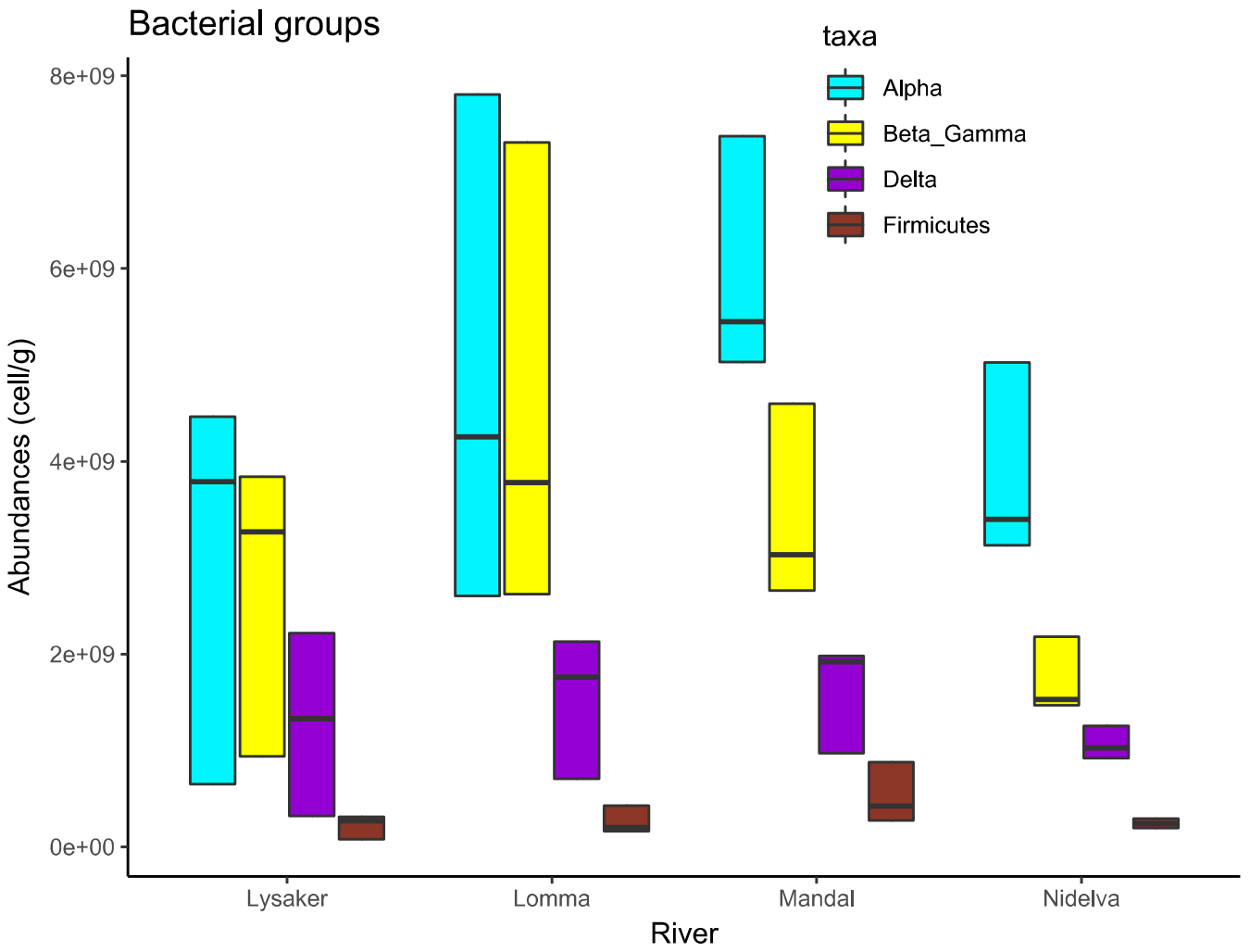


FIGURE 4 Comparison of cell numbers for the heterotrophic bacterial taxa analysed by CARD-FISH in the four rivers impacted by dams and hydropower plants (Alphaproteobacteria, Beta- and Gammaproteobacteria [joint abundances], Deltaproteobacteria and Firmicutes). Different ratios between Alpha- and Beta_Gammaproteobacteria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

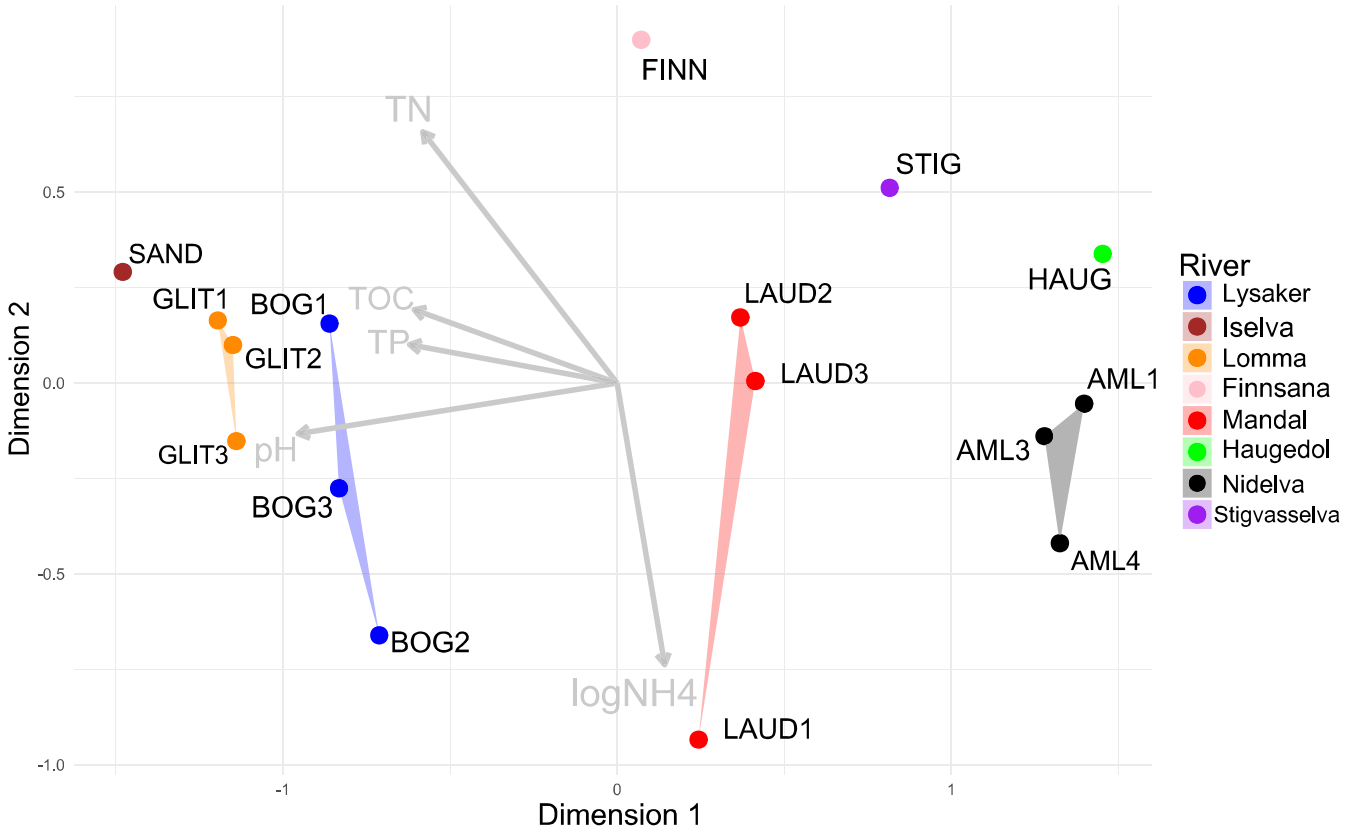


FIGURE 5 To analyse the distribution of the biofilm community structure and its relationship with the environmental parameters we plotted the envfit analysis produced by using the best GNMDS out of 100 iterations, performed on the Bray–Curtis matrix of ASV abundances and the data frame for the variables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

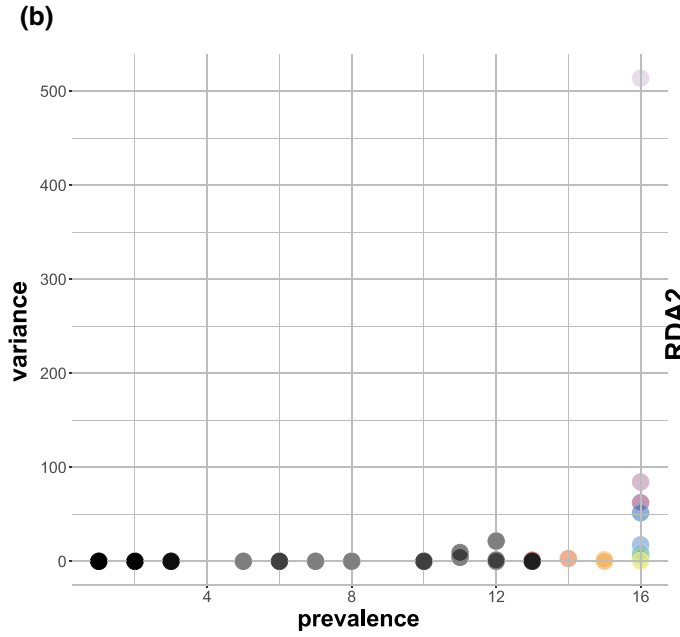
Betaproteobacteria were the third proteobacterial group to be highly represented in the sequencing data, with *Massilia* and *Janthiniobacterium* being the most common genera for this class. Betaproteobacteria are the group most associated with freshwater ecosystems, including important functional groups 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).

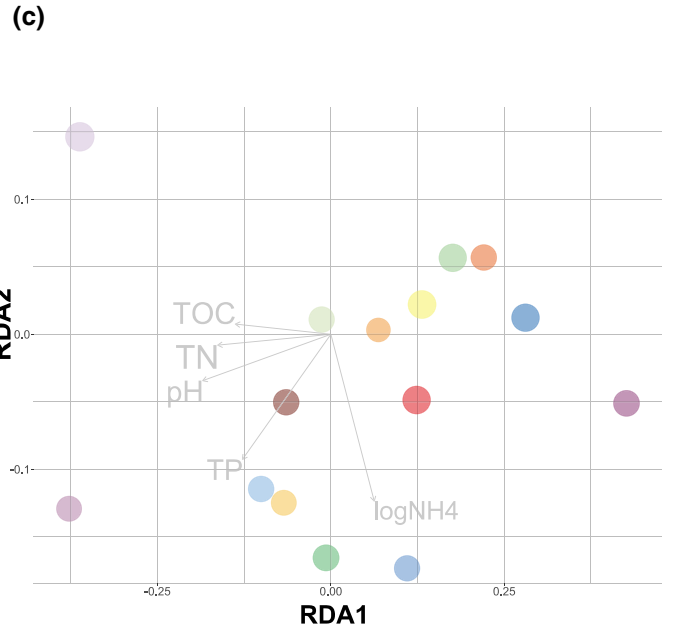
According to our findings, the microbial community structure is profoundly dependent on the physicochemical features of the region, confirming previous results on microbial communities from sediments in coastal areas, estuaries and rivers (Freixa et al., 2016; Aylagas et al., 2017; Borja, 2018; Fazi 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

MOLECULAR ECOLOGY RESOURCES

Supplemental Information for:

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

Lorenzo Pin, Alexander Eiler, Stefano Fazi, Nikolai Friberg

Table of Contents:

Table S1	Page 2
Table S2	Page 3
Table S3	Page 4
Table S4	Page 5
Table S5	Page 6
Table S6	Page 7
Figure S1	Page 8
Figure S2	Page 9
Supplementary information about the historical and geographic context of the analysed regions	Page 10
References	Page 11

MOLECULAR ECOLOGY RESOURCES

Table S1. Overview of regions, catchments, rivers and sampling sites.

Region	Catchment	River	Site	Treatment
South-eastern Norway (Oslo)	Lysakerelva	Lysaker	BOG1	control
			BOG2	impact
			BOG3	impact
	Sandvikselva	Iselva	SAND	control
		Lomma	GLIT1	control
			GLIT2	impact
			GLIT3	impact
South-western Norway (Agder)	Mandalselva	Finnsana	FINN	control
		Mandal	LAUD1	control
			LAUD2	impact
			LAUD3	impact
	Arendalvassdraget	Haugedol	HAUG	control
		Nidelva	AML1	control
			AML3	impact
			AML4	impact
		Stigvasselva	STIG	control

MOLECULAR ECOLOGY
RESOURCES

Table S2. Track table for the reads through every step of the bioinformatic processing.

Site	input	filtered	denoisedF	denoisedR	merged	nonchim
BOG1	44990	38558	36493	36914	34018	29866
BOG2	65324	54578	51769	52234	48592	48196
BOG3	64119	53095	50815	51297	48199	47303
SAND	62077	52687	51325	51511	49650	47869
GLIT1	54914	45658	44043	44460	41317	39250
GLIT2	60411	52494	50949	51347	47933	40771
GLIT3	57666	49149	46561	46930	42138	40086
FINN	47978	38407	36560	37062	34472	31901
LAUD1	52652	43147	40690	41328	37963	37502
LAUD2	56655	46909	45123	45453	43360	41907
LAUD3	60453	50419	48595	48981	46733	46204
HAUG	66571	53718	50241	51065	45438	45278
AML1	58067	46811	43327	44459	39798	39400
AML3	44523	35536	32518	33061	29072	28817
AML4	30324	16937	14165	15174	11650	11650
STIG	57629	47933	45275	45712	42413	41816

MOLECULAR ECOLOGY RESOURCES

Table S3. Chemical parameters analysed for the water samples from all the sampling sites.

Site	NH_4^+ (µg/l)	Ca^{2+} (mg/l)	Mg^{2+} (mg/l)	SO_4^{2-} (mg/l)	Fetot (mg/l)	DOC (mg/l)	TOC (mg/l)	TN (µg/l)	TP (µg/l)	Cond (µS/cm)	pH
BOG1	2	6.28	0.67	4.3	0.1	7	7.2	540	7	50.7	6.92
BOG2	29	6.38	0.74	4.13	0.12	7.2	7.3	570	12	53.5	7.12
BOG3	25	6.29	0.72	4.17	0.087	7.3	7.4	580	12	53.9	7.24
SAND	2	19.6	1.8	8.72	0.013	5	5	610	5	155	8.04
GLIT1	2	13.4	1.31	5.26	0.065	8.6	8.7	570	7	98.8	7.78
GLIT2	2	13.1	1.26	6.33	0.064	8.7	8.7	570	10	95.3	7.77
GLIT3	2	13.8	1.32	6.91	0.058	8.2	8.6	580	10	100.9	7.81
FINN	2	2.42	0.67	2.69	0.248	5.3	5.4	810	13	42.8	6.04
LAUD1	19	1.38	0.2	0.86	0.249	5.3	5.3	270	6	16.1	6
LAUD2	20	1.53	0.23	0.92	0.23	5.7	5.8	340	6	18.5	5.90
LAUD3	13	1.56	0.25	0.97	0.226	5.6	5.7	350	6	19.6	6.04
HAUG	4	0.55	0.17	0.97	0.338	7.5	7.5	280	5	15.5	4.80
AML1	11	0.88	0.16	0.97	0.0943	3.9	4	200	3	11.9	5.70
AML3	7	0.86	0.17	1.08	0.129	4.2	4.3	270	3	12.9	5.70
AML4	10	0.88	0.17	1.03	0.117	4.2	4.2	240	3	12.5	5.40
STIG	2	1.22	0.33	2.4	0.212	6.4	6.5	350	4	23	5.70

MOLECULAR ECOLOGY
RESOURCES

Table S4. Average abundances (cell/g) and standard deviation for total prokaryotes (DAPI), *Bacteria* and *Archaea*.

Site	DAPI	DAPI SD	Bacteria	Bacteria SD	Archaea	Archaea SD
BOG1	2,71E+09	3,86E+08	2,38E+09	54844377	1,39E+08	52368141
BOG2	1,64E+10	1,11E+09	1,41E+10	1,03E+09	7,29E+08	3,84E+08
BOG3	1,56E+10	3,55E+08	1,38E+10	5,61E+08	7,34E+08	5,08E+08
SAND	3,99E+10	1,23E+09	3,6E+10	7,36E+08	1,76E+09	4,35E+08
GLIT1	7,7E+09	8,88E+08	6,62E+09	5,07E+08	4,45E+08	2,86E+08
GLIT2	2,1E+10	2,04E+09	1,87E+10	3,46E+08	1,25E+09	1,51E+08
GLIT3	1,39E+10	1,19E+09	1,2E+10	3,63E+08	6,83E+08	6,04E+08
FINN	2,66E+10	1,85E+09	2,36E+10	1,63E+08	1,62E+09	2,56E+08
LAUD1	2,53E+10	2,06E+09	2,15E+10	92323940	8,07E+08	2,8E+08
LAUD2	2,01E+10	2,15E+09	1,62E+10	1,49E+09	8,23E+08	6,78E+08
LAUD3	1,59E+10	1,03E+09	1,36E+10	3,56E+08	6,06E+08	4,79E+08
HAUG	2,42E+10	1,38E+09	1,96E+10	1,09E+09	7,82E+08	1,45E+08
AML1	1,26E+10	1,56E+09	9,74E+09	58013419	5,76E+08	1,92E+08
AML3	1,69E+10	2,67E+09	1,43E+10	1,01E+09	8,25E+08	2,89E+08
AML4	1,22E+10	1,05E+09	9,66E+09	1,3E+08	6,06E+08	2,34E+08
STIG	1,32E+10	4,78E+08	1,11E+10	3,61E+08	6,82E+08	3,65E+08

MOLECULAR ECOLOGY
RESOURCES

Table S5. Average abundances (cell/g) the bacterial groups analysed by CARD-FISH.

Site	Alpha	Beta_Gamma	Delta	Auto	Firmicutes	OthBac	OthDAPI
BOG1	6,48E+08	6,48E+08	3,21E+08	3,99E+08	7,76E+07	1,63E+06	1,87E+08
BOG2	4,46E+09	4,46E+09	2,22E+09	1,54E+09	3,09E+08	2,27E+09	1,59E+09
BOG3	3,79E+09	3,79E+09	1,33E+09	1,56E+09	2,68E+08	2,97E+09	1,13E+09
SAND	1,37E+10	1,37E+10	2,95E+09	1,40E+09	4,99E+08	1,17E+09	2,15E+09
GLIT1	2,60E+09	2,60E+09	7,07E+08	4,48E+08	2,00E+08	3,67E+07	6,37E+08
GLIT2	7,81E+09	7,81E+09	2,13E+09	8,11E+08	4,26E+08	2,15E+08	1,05E+09
GLIT3	4,25E+09	4,25E+09	1,76E+09	3,90E+08	1,61E+08	1,69E+09	1,17E+09
FINN	9,22E+09	9,22E+09	3,06E+09	1,15E+09	3,56E+08	3,76E+09	1,38E+09
LAUD1	7,37E+09	7,37E+09	1,98E+09	1,33E+09	8,77E+08	5,31E+09	3,03E+09
LAUD2	5,03E+09	5,03E+09	1,92E+09	8,52E+08	4,20E+08	4,95E+09	3,11E+09
LAUD3	5,45E+09	5,45E+09	9,73E+08	5,36E+08	2,74E+08	3,67E+09	1,68E+09
HAUG	5,31E+09	5,31E+09	2,01E+09	5,01E+08	2,91E+08	7,04E+09	3,83E+09
AML1	3,40E+09	3,40E+09	9,21E+08	3,69E+08	2,42E+08	3,34E+09	2,28E+09
AML3	5,03E+09	5,03E+09	1,26E+09	9,24E+08	2,91E+08	4,61E+09	1,77E+09
AML4	3,13E+09	3,13E+09	1,02E+09	3,54E+08	1,92E+08	3,43E+09	1,95E+09
STIG	4,83E+09	4,83E+09	1,92E+09	2,77E+08	1,77E+08	7,67E+08	1,46E+09

MOLECULAR ECOLOGY
RESOURCES

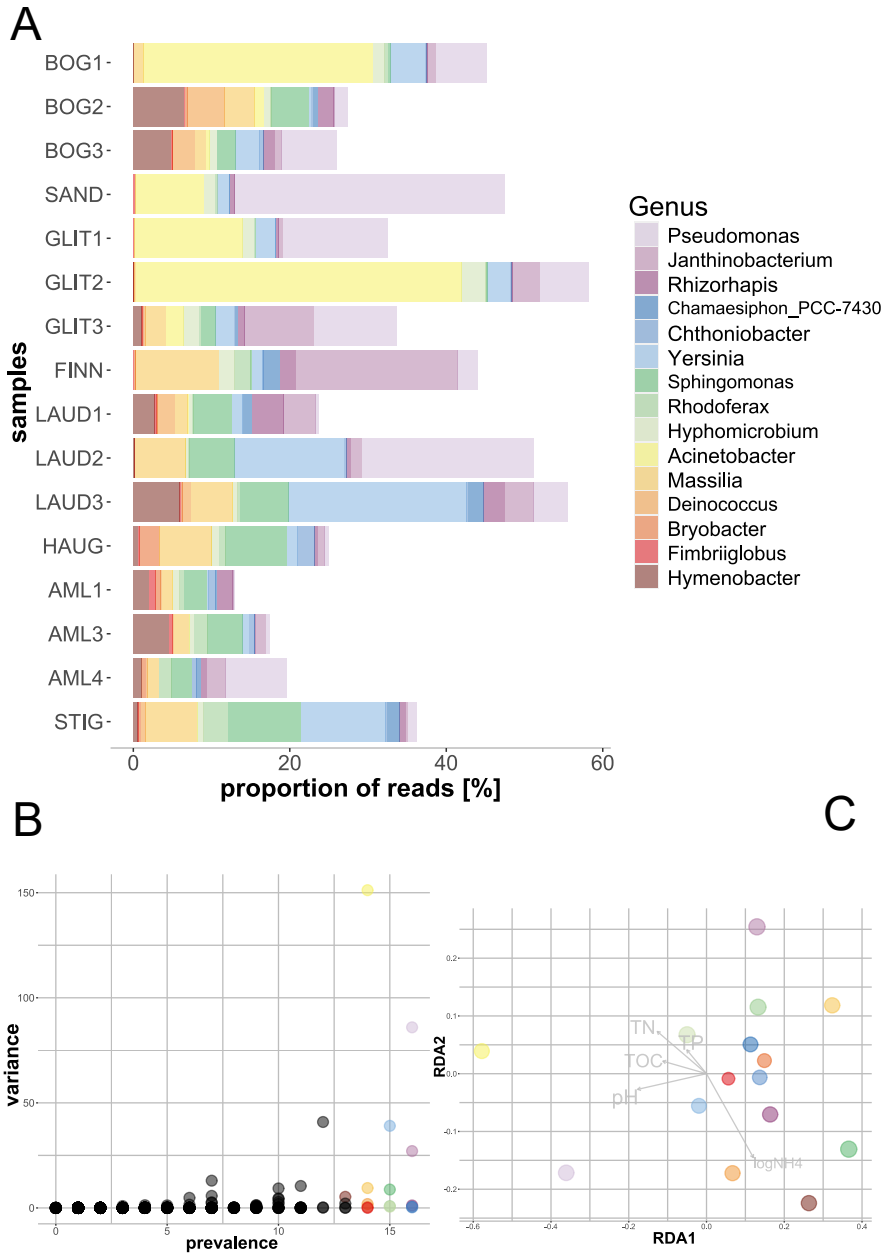
Table S6. Results from the envfit analysis did on the ASVs abundances and the environmental parameters.

VECTORS				
	NMDS1	**NMDS2**	**r2**	**Pr(>r)**
TOC	**-0.96588**	**0.25897**	**0.4269**	**0.022 ***
TN	**-0.66715**	**0.74492**	**0.7887**	**0.001 *****
TP	**-0.99158**	**0.12946**	**0.4163**	**0.024 ***
pH	**-0.98847**	**-0.15141**	**0.9148**	**0.001 *****
logNH4	**0.18951**	**-0.98188**	**0.5814**	**0.004 ** **

Signif. codes: 0 '*' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1**				
Permutation: free				
Number of permutations: 999				

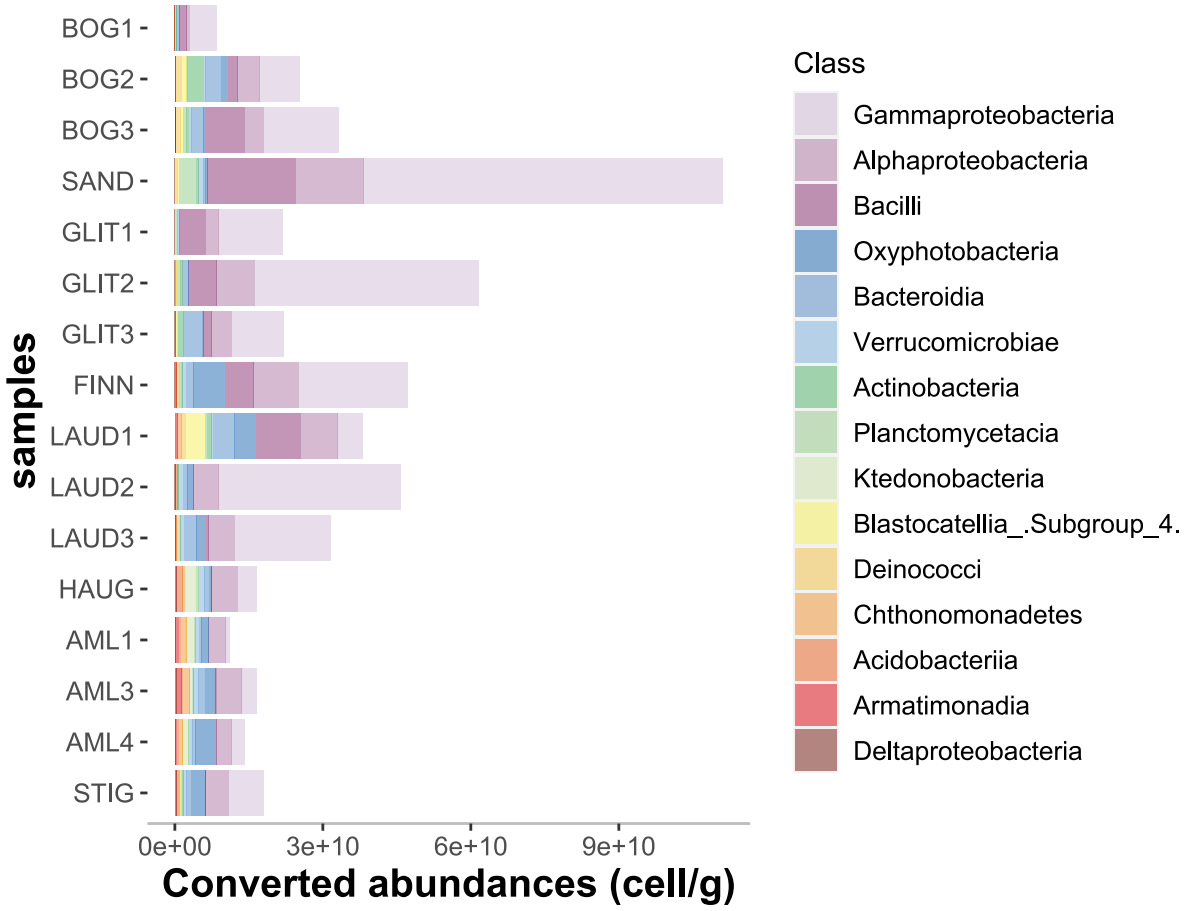
MOLECULAR ECOLOGY
RESOURCES

Figure S1. A) Barplot for the 15 most prevalent *genera* obtained by metabarcoding. B) Scatter plot for the prevalence and variance of the prokaryotic *genera*, highlighted with colors the 15 *taxa* showing high prevalence and high variance. C) RDA showing the distribution patterns for the 15 *genera* with the highest prevalence and highest variance and the environmental drivers. This approach would be suitable to detect prokaryotic bioindicators and set thresholds for water quality parameters. This graph provides an estimate for the most suitable *taxa* to be used as biological indicators given their broad presence and wide variance across different environments.



MOLECULAR ECOLOGY
RESOURCES

Figure S2. Absolut abundances for classes from metabarcoding were obtained from a converting factor based upon *Alpha-Proteobacteria* abundances for CARD-FISH. We divided each site's *Alpha-* abundance (CARD-FISH) by the relative abundance (metabarcoding results) for *Alpha-* in the same site, obtaining 1% absolute abundance for each site, which was then multiplicated by the percentages for each bacterial *class* within each site. Shown the 15 most abundant *taxa*.



MOLECULAR ECOLOGY
RESOURCES

Supplementary information about the historical and geographic context of the analysed regions

The discrepancy in the geology of the two regions is also reflected in various water chemistry parameters. The rivers in the Oslo region have, generally, higher conductivity, nutrients and higher pH due to the bedrock characteristics, the heavy agricultural utilisation of the areas surrounding Oslo and the urban and suburban developments which make this area the most densely populated of Norway (Reimann et al., 2009). On the other hand, the rivers in Agder, are generally low in nutrients and ions, due to the volcanic origin of the bedrock and the absence of sedimentary rocks. There is also less agriculture and lower population density in this area, compared to the Oslo region, indicating lower nutrient loads (Skjelkvåle et al., 2007).
During the seventies, the Northern regions of the world suffered from acidic precipitations, leading to acidification of water bodies in these regions (Stoddard et al., 1999). Ironically, after slowly recovering from the acidification issues, the rivers in south-western Norway are now suffering for calcium depletion, due to the reduced mobilization of the ion as a result of reduced deposition of H_2SO_4 (Hessen et al., 2017).The entire area around Oslo instead shows several anthropogenic impacts, which are heavily
impairing freshwater ecosystems including the rivers object of our investigation (Allan & Ranneklev, 2011; Grøndahl-Rosado et al., 2014; Nizzetto et al., 2016; Eregno et al., 2018).

References

Allan, I. J., & Ranneklev, S. B. (2011). Occurrence of PAHs and PCBs in the Alna River, Oslo (Norway). Journal of Environmental Monitoring, 13(9), 2420. https://doi.org/10.1039/c1em10314h

Eregno, F. E., Tryland, I., Tjomsland, T., Kempa, M., & Heistad, A. (2018). Hydrodynamic modelling of recreational water quality using Escherichia coli as an indicator of microbial contamination. Journal of Hydrology, 561, 179–186. https://doi.org/10.1016/j.jhydrol.2018.04.006

Grøndahl-Rosado, R. C., Tryland, I., Myrmel, M., Aanes, K. J., & Robertson, L. J. (2014). Detection of Microbial Pathogens and Indicators in Sewage Effluent and River Water During the Temporary Interruption of a Wastewater Treatment Plant. Water Quality, Exposure and Health, 6(3), 155–159. https://doi.org/10.1007/s12403-014-0121-y

Hessen, D. O., Andersen, T., Tominaga, K., & Finstad, A. G. (2017). When soft waters becomes softer; drivers of critically low levels of Ca in Norwegian lakes. Limnology and Oceanography, 62(1), 289–298. https://doi.org/10.1002/lno.10394Hill, M. O. (1973). Diversity and Evenness: A Unifying Notation and Its Consequences. Ecology, 54(2), 427–432. https://doi.org/10.2307/1934352

Nizzetto, L., Butterfield, D., Futter, M., Lin, Y., Allan, I., & Larssen, T. (2016). Assessment of contaminant fate in catchments using a novel integrated hydrobiogeochemical-multimedia fate model. Science of The Total Environment, 544, 553–563. https://doi.org/10.1016/j.scitotenv.2015.11.087

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

Skjelkvåle, B. L., Borg, H., Hindar, A., & Wilander, A. (2007). Large scale patterns of chemical recovery in lakes in Norway and Sweden: Importance of seasalt episodes and changes in dissolved organic carbon. Applied Geochemistry, 22(6), 1174–1180. https://doi.org/10.1016/j.apgeochem.2007.03.040

Stoddard, J. L., Jeffries, D. S., Lükewille, A., Clair, T. A., Dillon, P. J., Driscoll, C. T., Forsius, M., Johannessen, M., Kahl, J. S., Kellogg, J. H., Kemp, A., Mannio, J., Monteith, D. T., Murdoch, P. S., Patrick, S., Rebsdorf, A., Skjelkvåle, B. L., Stainton, M. P., Traaen, T., Wilander, A. (1999). Regional trends in aquatic recovery from acidification in North America and Europe. Nature, 401(6753), 575–578. https://doi.org/10.1038/44114