

KU LEUVEN



Assembly processes of bacterioplankton communities in shallow lakes along anthropogenic gradients

Fabio Toshiro T. HANASHIRO

Supervisor: Prof. Dr. Luc De Meester
Co-supervisor: Prof. Dr. Caroline Souffreau
Co-supervisor: Dr. Shinjini Muhkerjee

Chair of the Examination Committee:
Prof. Dr. Koenraad Muylaert

Members of the Examination Committee:
Prof. Dr. Eva Lindström
Prof. Dr. Bart Lievens
Prof. Dr. Luc Brendonck
Prof. Dr. Dirk Springael
Prof. Dr. Pieter Lemmens

Dissertation presented in partial fulfilment of the requirements for the degree of Doctor of Science (PhD): Biology

August 2022

© 2022 FABIO TOSHIRO T. HANASHIRO

Uitgegeven in eigen beheer, FABIO TOSHIRO T. HANASHIRO, LEUVEN

Alle rechten voorbehouden. Niets uit deze uitgave mag worden vermenigvuldigd en/of openbaar gemaakt worden door middel van druk, fotokopie, microfilm, elektronisch of op welke andere wijze ook zonder voorafgaandelijke schriftelijke toestemming van de uitgever.

All rights reserved. No part of the publication may be reproduced in any form by print, photoprint, microfilm, electronic or any other means without written permission from the publisher.

ACKNOWLEDGEMENTS

I would like to thank Luc and Caroline for the opportunity, help and patience during my PhD. Many thanks for the guidance and for sharing your knowledge with me along these years. My appreciation extends to Shinjini as well.

I am grateful to my funding agency CNPq for the opportunity to do a PhD in Belgium.

I really appreciated all the comments, suggestions and help from my Examination Committee (Prof. Dr. Koenraad Muylaert, Prof. Dr. Eva Lindström, Prof. Dr. Bart Lievens, Prof. Dr. Luc Brendonck, Prof. Dr. Dirk Springael and Prof. Dr. Pieter Lemmens). Thank you very much to all of you and specially to Eva – your work has been an inspiration to me since the beginning of my PhD.

Many thanks to all the collaborators of the different projects that I had the opportunity to work with during these years.

This was a long journey, and probably the best part of, it was the chance to meet and work with some incredible people at KU Leuven. I have been part of different "generations" of the De Meester's group and the Ecology, Evolution & Biodiversity Conservation Division and lucky enough to call some of you, my friends. Thanks! For all the kindness and logistic help, special thanks to Ria and Conny.

All my love and gratitude to my friends in Leuven. Without you guys, this would not be possible: Bubble! (Leticia, Susanna, Ricardo, Giancarlo and Gabriel – well, we survived a pandemic together…I think we can manage some distance, right?); De Meester's group (Matthias, Edwin, Laura, Jessie, Koen, Carla, Io, Haile, Jelena, Eyerusalem, Fran, Lorena, Lei, Andros, Leticia, Katina, Kristien, Laurens and Mariuxi – a special mention to Lynn, Rafaela, Konrad and Héléne, you guys were very important to me towards the end of this journey); Sem/Com Panelerem (Vanessa, Gui, Fer Machado, Fer Sodré, João, Fanny, William, Karla, Nadia, Camila, Gabi, Adrian, Els, Ana, Cadu, Cassio, Pedro, Oscar, Glauco, Rosalia, Guto, Felipe, André, Marcão, Eduarda and Tati); LSSE (Piet, Nadina, Jelle, Sascha, Tom, Beatriz, Eliseo, Annette and Ganso) + Korcan, Mireia, Magdalena, Nellie, Sylwia, Fabião, Franz, Hendrik, Breno, Victor, Tiziana and Sophie. Thanks a lot, to all of you for the friendship, good memories and support.

I would like also to thank my Brazilian friends that even from distance are somehow part of this work. You are always with me: Bio-03, Erto?, my "rep" (Tavin, André and Will), DEBE, Partido, Bio-UFSCar, Bundões and G.V.

To be away from my family, specially during the tough times was probably the most difficult part of living abroad. That's why I am forever grateful to my family for the love, support and understanding along these years. All my gratitude to my parents, my sister Jú, to the Los Taquicawas :), to the primos Filó and to Cintia's family.

Ci, you are the love of my life. Thank you for everything. Love, understanding and patience, I couldn't ask for more. You are the main reason why this thesis has finally come to an end. You are an incredible woman and researcher that I am so proud to be together.

I dedicate this thesis to my batiam Tereza and to my tio Massa.

SUMMARY

Bacteria play a fundamental role in the most important biogeochemical cycles in terrestrial and aquatic ecosystems. Despite the importance of bacterial communities for ecosystem processes and biodiversity, microbial ecology was for a long time limited to cultivation for the identification of species. However, over the last decades, a revolution occurred due to Next Generation Sequencing (NGS) methods that allow the identification of an elevated number of species and the possibility to correlate functional roles to the different species. As other communities, bacteria living in freshwater systems are facing different pressures related to anthropogenic actions (e.g., pollution, land use change and climate change). Thus, understanding how these bacterial communities are dealing with these different stressors and anthropogenic changes is urgent. Here, using the metacommunity framework, we evaluated the effects of different environmental pressures on bacterioplankton assembly. We estimated the relative importance of local and regional processes on taxonomic and functional profiles along gradients of environmental change. We carried out an extensive field sampling campaign along gradients of urbanization and eutrophication covering in total 145 shallow lakes and ponds. Besides, we performed a mesocosm experiment with 80 mesocosms to evaluate under experimental conditions how bacterioplankton assembly is affected by different levels of nutrient-enrichment, temperature increase and a zooplankton richness gradient. In chapter II, related to an urbanization gradient, we reported changes in bacterioplankton community composition in urban areas related to local environmental conditions, mainly driven by pH. This suggests species sorting as the prevalent assembly process in an urbanization context. In chapter III, we showed that local environmental conditions mediated by phytoplankton and zooplankton were more important in driving bacterioplankton assembly along a eutrophication gradient compared to regional factors. This result indicates the importance to include biotic interactions as explanatory variables in studies of bacterioplankton assembly. In chapter IV, we combined taxonomic composition, functional group annotations and carbon utilization profiles to compare the local and regional drivers in bacterioplankton assembly. Local environment was the main driver for functional group annotations and taxonomic composition. Carbon utilization profiles could not be explained by environment or space. This pattern suggests that functional traits are complementary to understand bacterioplankton assembly, especially because functional

traits were driven by different local variables. In chapter V, we used mesocosms under a setting of nutrient-enrichment, temperature increase and zooplankton richness to reproduce eutrophication scenarios and evaluate how bacterioplankton assembly is affected by these conditions. Biotic interactions mediated by different zooplankton species impacted significantly heterotrophic bacteria assembly. Divergent communities at the final time point were found to be correlated to the presence of Cyanobacteria. Nutrient and temperature treatments were less important in the trajectory of bacterioplankton communities after 12 weeks. Overall, our combination of different approaches – observational (taxonomic, functional and inclusion of biotic factors) and experimental (mesocosms) – contributed to a better understanding of bacterioplankton assembly in the context of environmental change (urbanization and eutrophication).

SAMENVATTING

Bacteriën spelen een fundamentele rol in de belangrijkste biogeochemische cycli in terrestrische en aquatische ecosystemen. Ondanks het belang van bacteriegemeenschappen voor ecosysteemprocessen en biodiversiteit, bleef microbiële ecologie lange tijd beperkt tot de kweek van stammen voor de identificatie van soorten. In de afgelopen decennia heeft er echter een revolutie plaatsgevonden dankzij Next Generation Sequencing (NGS) - methoden die de identificatie van een verhoogd aantal soorten mogelijk maken en de mogelijkheid bieden om functionele rollen aan de verschillende soorten te correleren. Net als andere gemeenschappen worden bacteriën die in zoetwatersystemen leven geconfronteerd met verschillende vormen van druk die verband houden met antropogene acties (bijv. vervuiling, verandering in landgebruik en klimaatverandering). Het is daarom dringend noodzakelijk om te begrijpen hoe deze bacteriële gemeenschappen omgaan met deze verschillende stressoren en antropogene veranderingen. Hier hebben we, met behulp van het metacommunity-raamwerk, de effecten van verschillende omgevingsdruk op de assemblage van bacterioplankton geëvalueerd. We schatten het relatieve belang van lokale en regionale processen op taxonomische en functionele profielen langs gradiënten van milieuverandering. We voerden een uitgebreide veldbemonsteringscampagne uit langs gradiënten van verstedelijking en eutrofiëring die in totaal 145 ondiepe meren en vijvers besloegen. Daarnaast hebben we een mesokosmos-experiment uitgevoerd met 80 mesokosmen om onder experimentele omstandigheden te evalueren hoe de assemblage van bacterioplankton wordt beïnvloed door verschillende niveaus van nutriëntenverrijking, temperatuurstijging en een gradiënt van zoöplanktonrijkdom. In het eerste hoofdstuk, gerelateerd aan een verstedelijkingsgradiënt, rapporteerden we veranderingen in de samenstelling van bacterioplanktongemeenschappen in stedelijke gebieden gerelateerd aan lokale omgevingscondities, voornamelijk gedreven door pH. Dit suggereert dat *species sorting* het meest voorkomende assemblageproces is in een verstedelijkingscontext. In het tweede hoofdstuk toonden we aan dat lokale omgevingscondities, gemedieerd door fytoplankton en zoöplankton, belangrijker waren in het aandrijven van de structuur van bacterioplankton langs een eutrofiëringsgradiënt in vergelijking met regionale factoren. Dit resultaat benadrukt hoe belangrijk het is om biotische interacties op te nemen als verklarende variabelen in studies naar de assemblage van bacterioplankton. In het derde hoofdstuk hebben we taxonomische

samenstelling, functionele groepsannotaties en koolstofgebruiksprofielen gecombineerd om de lokale en regionale factoren in de assemblage van bacterioplankton te vergelijken. De lokale omgeving was de belangrijkste drijfveer voor functionele groepsannotaties en taxonomische samenstelling. Koolstofgebruiksprofielen konden niet worden verklaard door omgeving of ruimte. Dit patroon suggereert dat functionele eigenschappen complementair zijn om de assemblage van bacterioplankton te begrijpen, vooral omdat functionele eigenschappen werden aangestuurd door verschillende lokale variabelen. In het vierde hoofdstuk gebruikten we mesokosmen in een omgeving van nutriëntenverrijking, temperatuurstijging en zoöplanktonrijkdom om eutrofiëringsscenario's te reproduceren en te evalueren hoe de structuur van bacterioplankton door deze omstandigheden wordt beïnvloed. Biotische interacties gemedieerd door verschillende zoöplanktonsoorten hadden een significante invloed op de heterotrofe bacteriesamenstelling. Divergerende gemeenschappen op het laatste staalnamepunt bleken gecorreleerd te zijn met de aanwezigheid van Cyanobacteriën. Voedings- en temperatuurbehandelingen waren na 12 weken minder belangrijk in het traject van bacterioplanktongemeenschappen. Over het algemeen heeft onze combinatie van verschillende benaderingen – observationeel (taxonomisch, functioneel en inclusie van biotische factoren) en experimenteel (mesokosmen) – bijgedragen aan een beter begrip van de assemblage van bacterioplankton in de context van milieuverandering (verstedelijking en eutrofiëring).

TABLE OF CONTENTS

Chapter I - Introduction 1

1.1 Microbial communities in the context of global change 2

1.2 (Studying) drivers of bacterial community structure 5

1.3 Advances in the estimation of microbial diversity 11

1.4 Shallow lakes and ponds: ideal systems for metacommunity studies 14

1.5 Mesocosm experiments as intermediate approach between field surveys and laboratory experiments 15

1.6 Aims and thesis outline 15

Chapter II - Freshwater bacterioplankton metacommunity structure along urbanization gradients in Belgium 21

2.1 Abstract 22

2.2 Introduction 23

2.3 Materials and methods 26

2.4 Results 31

2.5 Discussion 38

2.6 Supplementary material for Chapter II 44

Chapter III - Bacterioplankton assembly along a eutrophication gradient is mainly structured by environmental filtering, including indirect effects of phytoplankton composition 55

3.1 Abstract 56

3.2 Introduction 57

3.3 Material and Methods 59

3.4 Results 62

3.5 Discussion 69

3.6 Supplementary material for Chapter III 75

Chapter IV - Environmental filtering drives both taxonomic and trait structure in bacterioplankton communities of shallow ponds 89

4.1 Abstract 90
4.1 Introduction 91
4.3 Methods 93
4.4 Results 99
4.5 Discussion 109
4.6 Supplementary material for Chapter IV 114
Chapter V - Bacterioplankton assembly is structured by zooplankton in a mesocosm experiment mimicking eutrophication and warming 123
5.1 Abstract 124
5.2 Introduction 125
5.3 Material and methods 127
5.4 Results 133
5.5 Discussion 142
5.6 Supplementary Material for Chapter V 148
Chapter VI – General discussion 177
6.1 General discussion 178
6.2 Impacts of anthropogenic gradients on freshwater bacterioplankton 179
6.3 Drivers of freshwater bacterioplankton metacommunities 181
6.4 Trait-based structure as a complement in the study of bacterioplankton communities 184
6.5 Sequencing methods and statistical consequences 185
6.6 Perspectives and future work 187
References 189

Chapter I – Introduction

1.1 Microbial communities in the context of global change

Humans have an extensive impact on the environment and its related biodiversity, both due to the rapid human population growth and human activities such as land use change, overexploitation of natural resources, pollution, introduction of invasive species and climate change (Tilman et al., 2017). Much of this impact is diffuse, for example, the increasing pollution levels worldwide (e.g., fertilizer and pesticide use in agriculture) or the induced changes in the global climate (Parmesan & Yohe, 2003; Moritz et al., 2008; Walther, 2010). Human activities and land use transformations also affect non-harvested natural populations in multiple, indirect and unintentional ways. For example, the removal of top predators in marine systems can have detrimental effects cascading along the trophic chain, leading to environmental changes and collapse of other fish species (Frank et al., 2005). These changes are detrimental for many species leading to a deep biodiversity crisis that ranges from genes to ecosystems (Barnosky et al., 2011). In a meta-analysis, Darimont et al. (2009) showed that average phenotypic changes in human-harvested systems (systems in which organisms are under strong anthropogenic pressures caused by fishers, hunters and plant harvesters) were much more rapid than changes in natural systems. However, human activities can also lead to land transformations that can be beneficial for some organisms. For example, the increase of nutrient levels in water bodies (known as eutrophication) due to agriculture practices can provide conditions to a rapid growth of specific green algae and Cyanobacteria species (Kosten et al., 2012; Huisman et al., 2018). Thus, for a subset of organisms, anthropogenic environments can be favourable as profitable surrogates of natural environments (Beneš et al., 2003), or as novel environments that provide suitable physical or biotic conditions (Hobbs et al., 2006; Møller, 2010).

Microorganisms may respond in a different way to anthropogenic pressures compared to macroscopic organisms due to their high capacity to cope with changing and even extreme conditions, such as high levels of salinity, temperature, and different types of pollution (Rothschild & Mancinelli, 2001; Tamburini et al., 2013; Martiny et al., 2015). Compared to macroorganisms, microorganisms are assumed to have specific characteristics that increase adaptation to the new environmental conditions driven by anthropogenic change (Saito & Koike, 2013). Traits such as a small size, fast generation time, high mobility (mainly passive) and large genetic variability seem to be important species traits that might favour their local

adaptability (McDonnell & Hahs, 2015). Microbial communities control many parts of the fundamental biogeochemical cycles in soil (Jansson & Hofmockel, 2020), marine (Buchan et al., 2014) and freshwater systems (Bai et al., 2012). Regarding their importance, there is a need to understand how these communities deal with recent and abrupt changes in their respective environments (Logares et al., 2013).

In this thesis, I specifically focus on freshwater bacterioplankton as study organisms. These bacteria are part of the plankton component that drift in the water column in freshwater systems. Like other aquatic microbial communities, freshwater bacterial communities play a fundamental role in biogeochemical processes, providing several ecosystem services. As the main responsible for the degradation and mineralization of organic compounds to their inorganic forms, freshwater bacteria are crucial components of the freshwater food web, having a bottom-up effect across the trophic chain (Newton et al., 2011). In lentic systems as lakes and ponds, bacterial communities are influenced by the natural hydrological dynamics of these systems such as seasonal mixing events and thermal stratification (Shade et al., 2012). Also, these lakes and ponds provide different habitats to bacteria e.g., along the water column, on macrophytes and hard surfaces and in the sediments, leading to different communities for the planktonic and benthic components of freshwater bacteria (Souffreau et al., 2018; Zeng et al., 2019). The dispersal of bacteria among these different habitats might be influenced by the morphometric characteristics of the lakes, for example with a greater resuspension from sediments in more shallow systems (Comte et al., 2017). Thus, the bacterioplankton drifting in the water column in these shallow systems might provide a representative sample of the bacterial diversity of these lentic water bodies.

Nowadays, two relevant and common threats to natural ecosystems and their microbial communities worldwide are the increase of urbanization in terms of land use and the eutrophication of freshwater systems across different spatial scales. Urban land area expansion dramatically increased in the last decades, transforming natural areas and consequently also the biological communities related (Grimm et al., 2008). The increase of urbanization is usually associated with a reduction of species diversity due to a range of different stresses that may be directly or indirectly related to urbanization (e.g., pollution, discharge of nutrients, temperature increase) (McKinney, 2006; Wang et al., 2011; Reese et al., 2016). Nevertheless, soil microbial communities sampled within semi-natural areas of cities showed high values of diversity (Reese et al., 2016), with similar numbers of distinct soil

microbial phylotypes as found in other biomes such as arctic, tropical and desert soils (Ramirez et al., 2014). In aquatic systems, studies have shown that urbanization affects bacterial composition rather than alpha diversity. Species turnover related to organisms more or less adapted to urban conditions were reported in ponds (Hu et al., 2018) and streams (Wang et al., 2011; Simonin et al., 2019) along urbanization gradients. Hence, it seems that microbial communities may respond differently to urbanization depending on the habitat type.

Eutrophication is becoming a more frequent and significant threat to freshwater systems (Scheffer et al., 2001; Vilar et al., 2014). Nutrient pollution related to runoff from agriculture areas and warmer temperatures due to climate change create favourable conditions for the dominance of specific green algae and Cyanobacteria species that can reach high numbers in a short period of time (Paerl & Huisman, 2008). Direct changes in shallow lakes due to eutrophication include the increase in primary productivity, but also increase in turbidity due to higher phytoplankton densities, while indirect consequences are the release of toxic compounds related to Cyanobacteria blooms which can have detrimental effects on the freshwater bodies and their communities (Scheffer et al., 2001; Horner-Devine et al., 2003; Huisman et al., 2018). However, the consequences of eutrophication on bacterioplankton communities could be different, with some species favoured by the eutrophic conditions and other species filtered out. In the context of climate change, it is expected that global warming could benefit some species related to eutrophication. Certain Cyanobacteria species, for example, grow better at higher temperatures compared to other phytoplankton species. Additionally, warmer water temperatures are related to increased lake stratification, reducing water column mixing (Paerl & Huisman, 2008). With increased temperatures, algae blooms could become more frequent and less seasonal, even in temperate zones due to the effect of warmer temperatures on strengthening the stratification of the water column, reducing vertical mixing, favouring Cyanobacteria to access light outcompeting other phytoplankton species (Huisman et al., 2018). The effects on the freshwater bodies could be dramatic, affecting their biodiversity and consequently the ecological processes related (Scheffer et al., 2001; Huisman et al., 2018). Since microbial communities play a main role in all global biogeochemical cycles, it is pivotal to improve our knowledge on how these organisms are responding to eutrophication and urbanization to maintain biodiversity and ecosystem services.

1.2 (Studying) drivers of bacterial community structure

The metacommunity framework: integrating local and regional processes

For a long period, the study of microbial community assembly was defined by the "Everything is everywhere, but the environment selects" hypothesis, which assumes that microscopic organisms have no dispersal limitation and that only local conditions are important in structuring their communities (De Wit & Bouvier, 2006; Lindström & Langenheder, 2012). However, with advances in microbial ecology (e.g., identification due to sequencing methods and culturing) and the increased number of studies on microbial community assembly, it became clear that microbes can be dispersal limited, even at small spatial scales (Fig, 1A). Evidence of microbial biogeographic patterns with both local (environmental conditions) and regional (spatial effect) being important has been reported across different microbial groups, habitat types and spatial scales (Fig. 1A).

The study of the relative importance of local and regional factors for the assembly of communities was summarized into the metacommunity framework (Leibold et al., 2004). The framework identifies four different perspectives for metacommunities (local communities that interact by dispersal): (i) patch dynamics – assumes identical habitat patches and community composition is determined by the trade-off between species' competitive and colonization abilities (dispersal) with a species being a good disperser but bad competitor or the way around; (ii) species sorting – assumes that dispersal is relatively high and local environmental conditions determine community composition; (iii) mass-effects – assumes massive immigration of species that can be rescued in local habitats where they are bad competitors, diminishing the importance of niche requirements and (iv) neutrality – assumes that all species in the regional pool are identical in their competitiveness and dispersal abilities resulting in community composition being driven by drift. Local processes encompass primarily niche-related mechanisms related to environmental selection (more known as species sorting in metacommunity ecology) such as abiotic conditions (Soininen, 2014) or biotic interactions (Souffreau et al., 2015). Mass effects, dispersal limitation and neutral assembly are examples of regional processes assembling bacterioplankton communities (Leibold et al., 2004; Lindström & Langenheder, 2012). Dispersal between suitable habitat patches depend on the dispersal capacity of organisms and their regional abundance, while the identity of immigrants depends amongst others on the configuration of habitat quality in

space. Spatial configuration of habitat patches may strongly influence how the environment is tracked by populations and communities (Kozak et al., 2008). Microorganisms such as bacteria can easily disperse due to their high abundance and very small size, which makes them carried by the wind and other vectors (water, animals and plants) (Mony et al., 2020). Therefore, bacterial community composition within a metacommunity could increase in dissimilarity depending on the distance among habitat patches (Bell, 2010).

The metacommunity ecology approach has been used and developed in the past decades and has provided valuable insights in how communities are assembled and impacted by human activities (Logue et al., 2011). Organisms widely differ in reproductive mode and generation time, and this strongly impacts the strength of local responses. More specifically, microorganisms tend to be efficient in tracking environmental change, even at large geographic scales due their short generation times and rapid population growth rates (Van Der Gucht et al., 2007). A field study on small standing water bodies (ponds and shallow lakes) sampling 11 different organism groups showed that with increasing organism size the impact of environmental selection decreases and the impact of spatial isolation on species composition increases (De Bie et al., 2012). Community assembly of small organisms, however, may also be partly determined by processes governing communities at higher trophic levels (Verreydt et al., 2012). For example, bacterioplankton communities in shallow lakes along a latitudinal gradient are mainly driven by water chemistry and grazer abundance (pelagic cladocerans and cyclopoid copepods) (Souffreau et al., 2015). This makes these organisms especially interesting to study how anthropogenic impacts are changing community assembly.

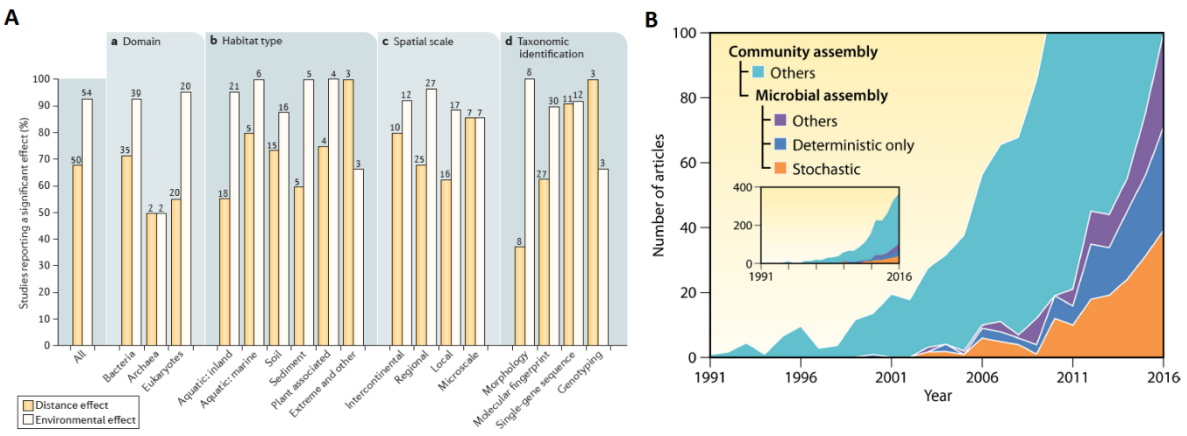

Figure 1. (A) Overview of the number of articles reporting significant effects of distance (space) and environment (local variables) underlying microbial assembly in different contexts and using different approaches. Retired from Hanson et al. (2012). (B) Number of articles studying community assembly for different organisms (light blue) and for microbial assembly specifically showing the roles of different mechanisms. Retired from Zhou & Ning (2017).

It is important to note that local and regional processes are not necessarily mutually exclusive and an interplay between both is probably the general rule (Langenheder & Székely, 2011; Stegen et al., 2013; Zhou & Ning, 2017; Langenheder & Lindström, 2019) with their contribution being dependent of other processes such as legacy effects (Andersson et al., 2014), rarity (Liu et al., 2015) and habitat specialization (Székely & Langenheder, 2014). Nevertheless, the relative importance of local and regional processes on the assembly of biological communities might be modified due to environmental changes (Langenheder & Lindström, 2019). Whereas environmental change could lead to a reinforcement of environmental selection (e.g., acidification and desiccation) (Valverde et al., 2014; Tripathi et al., 2018), other studies have shown this also can lead to an increase in stochasticity. In a large-scale observational study in Scotland, Powell et al. (2015) showed that regional processes were more important assembling soil communities under more intensive land-use regimes (e.g., arable land). Using an experimental approach, Zhang et al. (2016) showed that nutrient-enriched conditions increased the relative importance of regional over local processes on soil communities. Similarly, bacterioplankton community assembly in mesocosms was mainly mediated by regional processes when submitted to a combination of fertilization and warming (Ren et al., 2017).

The importance of biotic interactions in structuring bacterial communities

Interactions among different biotic components within the metacommunity are becoming recognized as important factors structuring microbial communities. Due to the key role of bacterial communities in biogeochemical cycles (Buchan et al., 2014), possible changes in their biotic interactions might have an effect throughout the food chain impacting related ecological functions (Fig. 2). Top-down effects based on predation are probably the most described drivers inducing changes in bacterioplankton diversity and composition. Zooplankton, and ciliated or flagellated protists can favour or suppress specific bacterioplankton species (Pernthaler, 2005). An observational study assessing the relative

importance of local and regional processes along a latitudinal gradient (Souffreau et al., 2015) has shown that densities of the grazer *Daphnia* could explain a significant part of bacterioplankton variation. Manipulation experiments carried out in *Daphnia*-dominated ponds suggested a strong predation effect on the structure of bacterial communities, and thus not only in terms of bacterial abundances and biomass as previously expected (Langenheder & Jürgens, 2001). Such changes in bacterial community composition due to predation might be the result of predator-defence mechanisms (e.g., exopolymer formation, motility traits, changes in cell wall structure and toxin release) or size-selective grazing by the predators which might again be counteracted by the prey by cell size reduction and filament formation (Pernthaler, 2005).

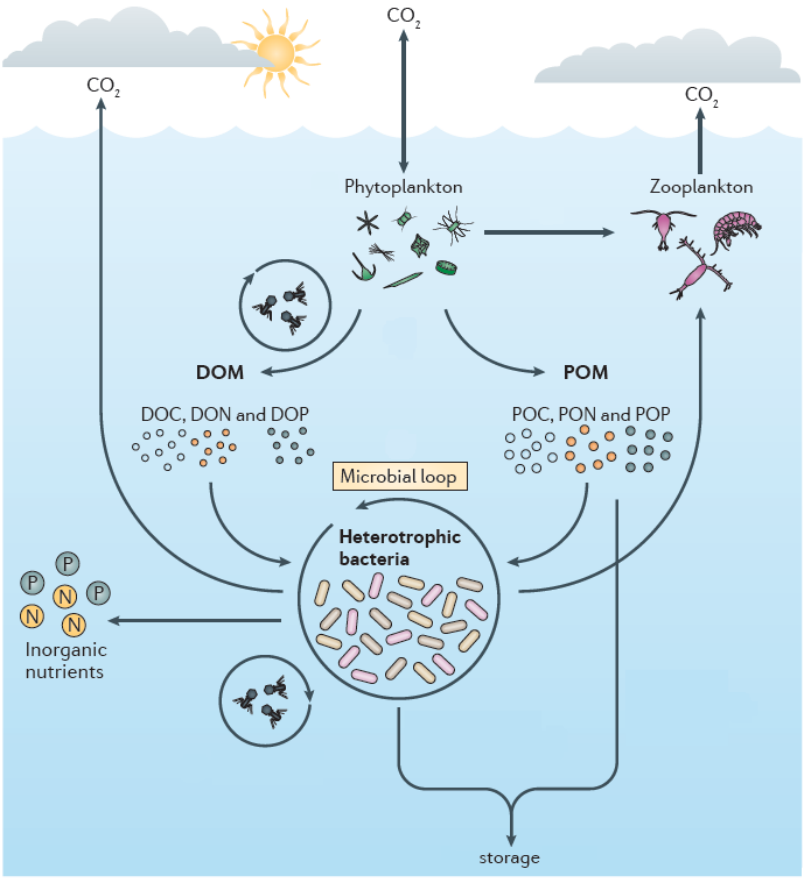

Figure 2. Typical food web in aquatic ecosystems highlighting the key role of microbes in biogeochemical cycles. Abbreviations are related to dissolved (e.g., DOC – dissolved organic carbon) and particulate (POP – particulate organic phosphorus) components. Modified from Buchan et al. (2014).

Viral lysis is also an important biotic interaction in bacterial communities and it is considered one of the main drivers of bacterial mortality in aquatic systems (Zimmerman et al., 2020). Viruses might be responsible in regulating the proportions of the most common

prokaryotes in freshwater lakes (Ram et al., 2014). This viral effect alters nutrient cycling in the microbial loop through the viral shunt pathway, "shortening" the availability of dissolved organic matter (DOM) through the lysis of heterotrophic bacteria and phytoplankton (Zimmerman et al., 2020). Phytoplankton and Cyanobacteria also significantly affect heterotrophic bacteria in aquatic systems over localized and seasonal increases in abundance known as algae blooms. Due to the increase in availability of organic matter during these algae blooms, a subsequent increase in the activity of heterotrophic bacteria has been reported (Buchan et al., 2014). Effects on heterotrophic communities related to algae and Cyanobacteria blooms were also described in relation to different ecological aspects such as bacterial mobility (e.g., free-living and particle-associated bacteria having different responses to the algae and Cyanobacteria growth) (Parveen et al., 2011; Xu et al., 2020), *Microcystis* presence (Su et al., 2017) and temporal changes – variable responses of heterotrophic bacteria in short period of times (Woodhouse et al., 2016). But even with this extensive list of examples in which biotic interactions might be important to impact microbial composition and diversity, the inclusion of biotic factors as explanatory variables in bacterioplankton assembly studies is still limited (Langenheder & Lindström, 2019).

Functional traits as metrics to understand bacterioplankton assembly

A complementary approach in the study of community assembly that is becoming more frequently used is to incorporate functional data to the ecological analyses (Krause et al., 2014; Langenheder & Lindström, 2019). Instead of using only classical taxonomic or phylogenetic information in community assembly, trait-based ecology relies on the use of expressed functions, functional classifications or the composition of functional genes (Krause et al., 2014). The relative importance of local and regional factors in driving community assembly can be different when using taxonomic or functional classifications (Langenheder & Lindström, 2019). Additionally, different local environmental drivers can be important for the taxonomic and functional dimensions (Bissett et al., 2010; Ruiz-González et al., 2015b; Louca et al., 2016). For example, Ruiz-González et al. (2015) studying boreal freshwater ecosystems showed that trait structure was driven by a gradient of quality of dissolved organic matter (DOM), while taxonomic structure was controlled by other factors such as pH, water temperature and water residence time. It is also suggested that most of the operational

taxonomical units (OTUs) share similar functional traits, resulting in functional redundancy (Konopka, 2009; Comte & Del Giorgio, 2010).

Different methods have been developed to address the functional properties of individual microbial organisms (e.g., cell shape) to entire communities (e.g., respiration rates). Functional traits can be categorized in different classifications depending if the traits are measured based on the genotypic or phenotypic characteristics (Escalas et al., 2019). Genotypic traits methods rely on the quantification of the presence of genes or pathways related to specific functional processes [e.g., PiCRUSt – phylogenetic investigation of communities by reconstruction of unobserved states (Langille et al., 2013), paprica – pathway prediction by phylogenetic placement (Bowman & Ducklow, 2015) and metagenomic sequencing]. Another genotypic traits method recently created to estimate functional potential of bacterial taxa and communities is FAPROTAX – functional annotation of prokaryotic taxa (Louca et al., 2016). An advantage of FAPROTAX relies on the fact that the metabolic phenotypes are based on experimental evidence (e.g., culturing) found in literature. Thus, FAPROTAX may be considered as a realized ecological function rather than a potential metabolic pathway. FAPROTAX was used in chapter IV of the present thesis to estimate functional group annotations of bacterioplankton communities of the investigated shallow ponds.

Phenotypic traits methods can be considered as quantifications of the expression of genotypic traits and can be more related to the actual ecological processes. Examples of phenotypic traits include characteristics of organisms (e.g., cell shape, growth rate and stoichiometry) and measurements of biological properties (e.g., production of gases and substrate use) (Escalas et al., 2019). Metabolic community profiles are commonly used to functionally characterize microbial communities through the exposition of bacteria to fixed sets of carbon sources (Preston-Mafham et al., 2002). In chapter IV of this thesis, we used carbon utilization profiles from Biolog Ecoplates™. Each plate contains three replicates of 31 carbon sources of different types (amino acids, carbohydrates, carboxylic acids, amines and polymers) and three blanks. This method has been extensively used to estimate functional properties across different environments and contexts (Dickerson & Williams, 2014; Souffreau et al., 2014; Ruiz-González et al., 2015b; Rutgers et al., 2016). We acknowledge that the trait-based methods used in this thesis have some limitations. Carbon utilization profiles are growth dependent, since probably only part of the community actually consume the carbon sources.

FAPROTAX group annotation relies on a potential function, since we cannot know if the functional gene was switched on. However, we believe that these trait-based methods are useful for our exploratory approach to better understand on how bacterioplankton communities are responding to the environmental gradients.

Even in microbial communities presenting signs of dispersal limitation and historical contingency, the response in terms of functional structure could be redundant because of intrinsic microbial features as rapid adaptation and horizontal gene transfer (Green et al., 2008). Moreover, understanding the functional structure of microbial communities will provide insights to the bridge between organisms and ecosystem functions. Freshwater systems such as shallow ponds as well as their biological and functional attributes are threatened by multiple anthropogenic disturbances such as salinization (Cañedo-Argüelles et al., 2016), land use modification (Vörösmarty et al., 2010) and pollution (Yamamuro et al., 2019). Thus, understanding the functional characteristics of these communities is very relevant to understand how they are responding to environmental changes.

1.3 Advances in the estimation of microbial diversity

Microbes are of crucial importance for ecosystem functioning and ecosystem services provided by different ecosystems (Martiny et al., 2015). But it is only since that the link between these organisms and their ecological services could be experimentally investigated and described due to technical advances. Before, identification of microorganisms was limited since many of the bacterial taxa could not be cultured. Over the last two decades, molecular fingerprinting techniques and, more recently, next generation sequencing (e.g., Roche 454 pyrosequencing, Illumina Technology, PacBio and Oxford Nanopore) have made it possible to characterize taxon composition of local microbial communities through a metagenomics approach, where DNA is isolated from an environmental sample and characterized based on the relative frequencies of variants in DNA at specific loci (Pollock et al., 2018). These recent approaches allowed the study of microbial diversity in different environments, including marine and freshwater systems (Galand et al., 2009; Bai et al., 2012).

In the present thesis, we used two different methods of amplicon sequencing (454 pyrosequencing and Illumina MiSeq) to characterize bacterioplankton communities. The Roche 454 sequencer is no longer available, but until 2013-2014 it was the main platform used

for 16S rRNA gene metabarcoding studies. Nowadays, Illumina MiSeq became the most common sequencing platform used in metabarcoding studies (Pollock et al., 2018). Illumina MiSeq platform balances accurate reads and a higher throughput with a relatively low cost. Although these methods have significant differences in read length, sequencing protocols and clustering pipelines, both methods can provide reliable results (Luo et al., 2012; Sinclair et al., 2015). Other methods such as PacBio and Oxford Nanopore are increasing in use due to their capacity to sequence the full length of the 16S gene.

High-throughput sequencing platforms revolutionized microbial studies and made it possible to uncover the immense microbial diversity, including access to rare species. Microbial communities are often dominated by few species that occur in high proportions while most of the diversity is rare and occur in very low numbers (Pedrós-Alió, 2012). This pattern brings different consequences to the study of microbial communities, for example difficulties regarding determining alpha, beta and gamma diversity across communities. Different approaches have been used to overcome the disproportional number of rare species, such as analysing different categories of relative abundance separately (e.g., only the abundant ones with relative abundance higher than 1% across samples) (Baltar et al., 2015; Liu et al., 2015; Liao et al., 2017) or categorizing species by their habitat specialization (Székely & Langenheder, 2014).

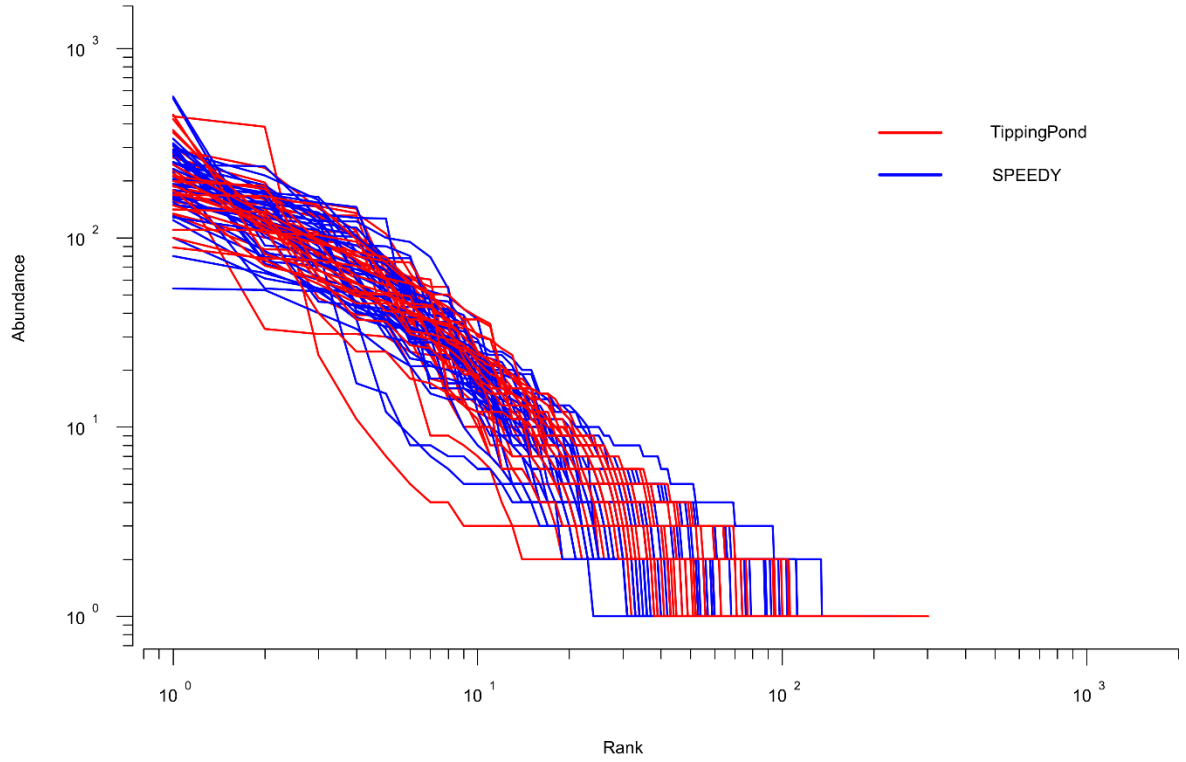

Figure 3. Typical rank-abundance (species are ranked based on their abundances – most abundant species = rank 1, the second most abundant species = rank 2 until the less abundant species = rank n, with n = total number of species) trend for microbial communities with few species/OTUs being abundant, while many OTUs are rare. Data from chapter IV (73 shallow ponds) – colours depict the Tipping Pond and SPEEDY sampled ponds.

Freshwater bacterioplankton from shallow lakes and ponds can also present this dominance of few species as it is demonstrated in the rank-abundance curve for the data of chapter IV (Fig. 3). While for alpha diversity, the inclusion of rare species gives a clearer perspective of how diverse microbial communities are, the addition of a high number of species with low abundances often introduce noise to the beta-diversity analyses. Different approaches have been proposed to improve the statistical analyses that could be affected by these issues. Most part of the community variation is often driven by the most abundant/common species, so one approach relies on the idea of running the analyses with different cut-off levels of relative abundance (≥ 0.1% and ≥ 1% for the abundant OTUs) or based in the number of reads per sample (Lynch & Neufeld, 2015). Another approach commonly used in beta-diversity studies that deal with a high number of rare species is to analyse community variation using distant-based analyses (e.g., PCoA - Principal Coordinate Analysis) - eigenvectors that synthetize community variation (Legendre & Andersson, 1999).

In this thesis, we only focus on the more dominant taxa, but it is important to notice that bacteria can have key roles in biogeochemical cycles even when present in low abundances. For example, the disproportional contribution of rare bacteria to the total carbon biomass in freshwater systems has been related to specific features such as large cell size (Šimek et al., 2014) and rapid growth (Neuenschwander et al., 2015).

1.4 Shallow lakes and ponds: ideal systems for metacommunity studies

Shallow lakes and ponds are well-studied systems known to support high levels of biodiversity of different semi-aquatic and aquatic organisms, despite their small size (Downing, 2010; Beklioğlu et al., 2016). Regardless of their importance in sheltering many species, shallow lakes and ponds are threatened by diverse anthropogenic pressures such as land use change, eutrophication, climate change and pollution (Declerck et al., 2006). In a metacommunity perspective, lakes and ponds are considered suitable systems to test hypotheses because they have clear boundaries between the different communities (De Meester et al., 2005). Besides, local environmental conditions of shallow lakes and ponds are influenced by the surrounding area (e.g., land use, shading and nutrients run-off) (Scheffer & Van Nes, 2007; Brans et al., 2018) making them also useful to evaluate changes across terrestrial-aquatic ecosystems. The spatial distribution of shallow lakes and ponds plays a key role in the way in which local communities can react to environmental change as this critically depends on the interaction between local and regional processes. Dispersal across different shallow lakes and ponds depends on the dispersal capacity of organisms, their regional abundance and the local environmental differences. Microbial organisms differ in reproductive mode, generation time and dispersal capacities among species and especially when compared to macroorganisms, and these different features could strongly impact the strength of their responses to local and regional processes. Different studies have illustrated how aquatic microbial communities can present biogeographical signatures depending on the local environmental conditions (Ruiz-González et al., 2015b), habitat specialization (Székely & Langenheder, 2014), spatial (Soininen et al., 2011) and temporal scales (Bell, 2010). Shallow lakes and ponds then might be

interesting systems to evaluate the possible effects of different stressors and perturbations on microbes under the metacommunity framework.

1.5 Mesocosm experiments as intermediate approach between field surveys and laboratory experiments

Mesocosms are small to middle-sized model systems that intent to mimic aquatic bodies and have been successfully used to test different hypotheses in metacommunity studies (Chase, 2010; Verreydt et al., 2012; Ren et al., 2017). Mesocosm experiments allow the control of different local conditions and can reproduce semi-natural conditions (Logue et al., 2011; De Meester & Pantel, 2014). Microbial studies could specially take profit of experimental set-ups because different mechanisms can be tested in relatively short temporal and small spatial scales (Konopka, 2009). Additionally, microbial studies in mesocosms tend to be less affected by “container effects” (changes in ecological responses due to captivity) because of their small size compared to the water volume of the container (Shade et al., 2018). Regardless of the advantages, results of this experimental approach should be interpreted with caution since mesocosm cannot fully capture the complexity of biological systems (see more in one of the classics "View from the Park" - Lawton 1999).

1.6 Aims and thesis outline

In this PhD thesis, we aimed to increase our understanding of the drivers of freshwater bacterioplankton community structure along anthropogenic gradients, with a specific emphasis on biotic factors (zooplankton) and including trait data. More specifically, we quantified the relative importance of local (environmental) and regional (spatial) factors on the taxonomic and functional community structure of bacterioplankton of shallow ponds over gradients of urbanization and eutrophication using field survey data and a mesocosm experiment.

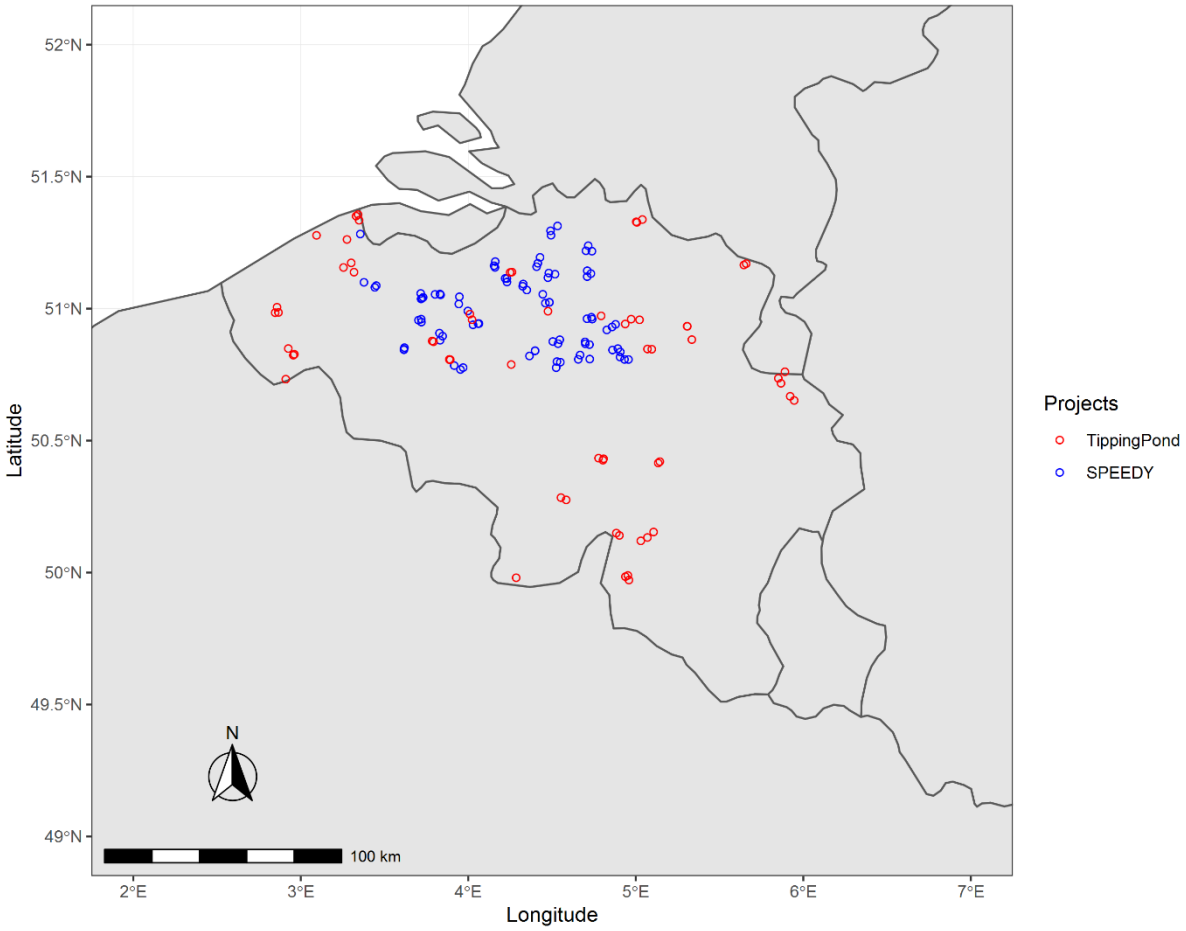

Figure 4. Map of Belgium with the location of the 145 ponds and shallow lakes sampled during this PhD thesis. Coloured circles represent the two different research projects (TippingPond and SPEEDY).

The thesis consists of four research chapters. A schematic overview of the thesis outline is represented in Figure 6. The first three chapters consist of observational studies while in the fourth chapter an experimental approach was used. In this thesis, we used data of bacterioplankton communities from 145 shallow lakes and ponds across Belgium, covering a sampled area of more than 20,000 km^2 with a focus on the Northern region (Flanders) (Fig. 4). This region is known for its heterogeneous landscape ranging from: (1) natural – forest patches and protected areas; (2) agricultural – intensive cultivation areas (e.g., crop fields), grasslands and (3) urban – cities and villages with different levels of population density and

built-up occupation areas. The sampled ponds were selected based on environmental gradients related to two different research projects. For the SPEEDY project, the selection of ponds was based on the percentage occupied by built-up area (e.g. buildings) [see more details in chapter II of this thesis and in Gianuca et al. (2018); Merckx et al. (2018)], while for the TippingPond research project, the selection was based on a land use gradient (Fig. 5) [see more details in chapter III of this thesis and in Declerck et al. (2006); De Bie et al. (2012); Verbeek et al. (2018)]. We sampled 145 shallow lakes and ponds in which initially 81 were used for the SPEEDY project (Chapter II) and 64 for the TippingPond project (Chapter III). Due to methodological issues, only a part of these samples could be used in Chapter II (50) and III (36). In Chapter IV, a combination of ponds (73) from the different projects were used, where 37 were from the SPEEDY dataset and 36 from the TippingPond dataset.

The aims of the four chapters can be summarized as follows:

In chapter II, we combined a metacommunity approach with a hierarchical sampling design to study the drivers of bacterioplankton community structure of shallow ponds along an urbanization gradient. Belgium is one of the most urbanized countries in North Europe, also very densely populated and with a high general human impact on land use. Specifically, we aimed to evaluate if urbanization led to changes in microbial diversity and community composition and to quantify the relative importance of local environmental conditions and spatial factors for bacterioplankton assembly. The hierarchical sampling design allowed also to disentangle effects of urbanization on two spatial scales (local and regional). We tested the hypotheses that (1) in urban landscapes, environmental factors are relatively stronger than spatial factors in regulating bacterioplankton community assembly; (2) urbanization leads to more homogeneous bacterioplankton community composition, with relatively minor effects on diversity.

In chapter III, we assessed whether incorporating biotic interactions (phytoplankton and zooplankton communities) improved the quantification of the relative importance of local conditions and regional processes in assembling bacterioplankton communities along a eutrophication gradient. Specifically, we combined three different aquatic communities (bacterioplankton, phytoplankton and zooplankton) sampled along a eutrophication gradient to evaluate if higher trophic levels directly or indirectly affected bacterioplankton community assembly through a path and a network analysis. Additionally, we estimated changes in bacterioplankton composition and diversity along the eutrophication gradient. We

hypothesized that (1) bacterioplankton assembly is mainly driven by environmental filtering in a eutrophication context, (2) If yes, biotic variables explain most of community variation compared to abiotic variables, and (3) eutrophication leads to higher bacterioplankton diversity.

In chapter IV, we assessed how the functional structure of bacterioplankton communities were shaped by local and regional factors using two trait-based methods (carbon utilization profiles and functional group annotations). Carbon utilization profiles assess the metabolic potential of bacteria to consume specific carbon sources, whereas functional group annotations categorize the identified taxa into functional groups (e.g., chemoheterotrophs) based on a taxonomic database. These two functional trait dimensions could help us to understand how these organisms perceive the surrounding environment. We then investigated if the different dimensions (taxonomic or functional – in terms of carbon consumption and functional group annotations) were structured by the same or by a different set of factors and which factors were important for each dimension. In summary, we hypothesized that (1) functional traits are strongly driven by environmental factors; and (2) a different set of variables influence the different community dimensions.



Figure 5. Mesocosm experiment at Arena site at KU Leuven campus (photo credit: Matthias Vanhamel).

In chapter V, we used an experimental set-up of 80 mesocosms simulating shallow lakes and ponds to evaluate the effects of nutrient-enrichment, warming and zooplankton richness on bacterioplankton assembly. We inoculated a similar bacterioplankton community to all mesocosms and after a twelve-week period evaluated changes in community composition and diversity across the different treatments. We hypothesized that (1) fertilization and warming lead to more productive environments, allowing different bacterial

species to flourish in these conditions, driving changes in community structure; and (2) nutrient-enrichment leads to algae blooms and to eutrophication scenarios in which the importance of zooplankton mediating community dynamics affects bacterioplankton assembly. We specifically choose mesocosms because these allow us to simulate the complexity of shallow lakes and ponds while being easy to manipulate the specific conditions. Eighty 210 L tanks were set at Arena area at KU Leuven campus (Fig. 5) and were maintained for 12 weeks under different combinations of nutrient-enrichment, warming, zooplankton richness, and presence of the keystone freshwater species *Daphnia magna*.

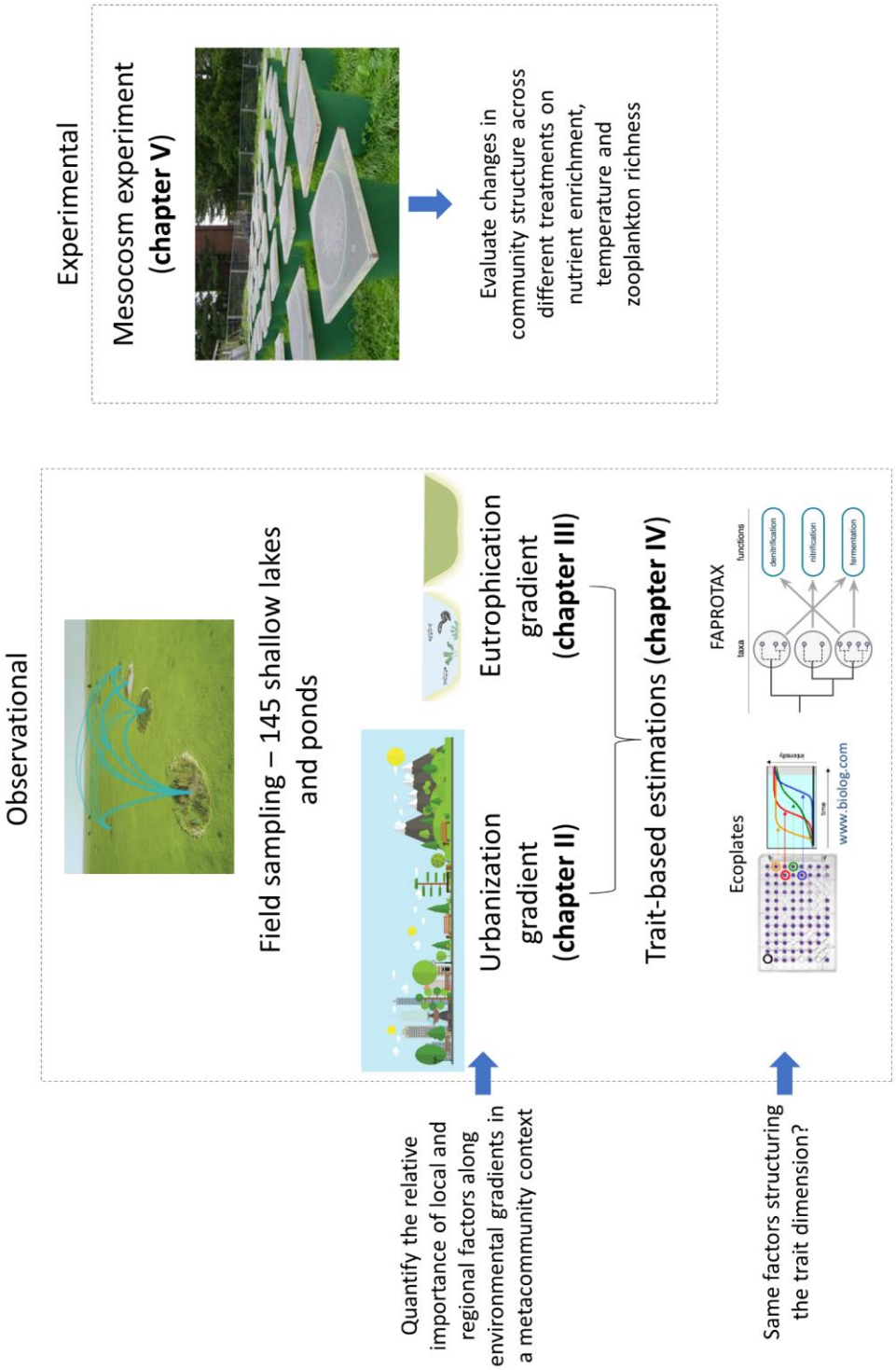

Figure 6. Schematic overview and integration of the different chapters of this thesis. Figures adapted from: bio.kuleuven.be/eeb/ldm/research; vecteezy.com; biolog.com; loucalab.com.

Chapter II

Freshwater bacterioplankton metacommunity structure along urbanization gradients in Belgium

Fabio Toshiro T. Hanashiro[1], Shinjini Mukherjee[2], Caroline Souffreau[1], Jessie Engelen[1], Kristien I. Brans[1], Pieter Busschaert[3,4] & Luc De Meester[1]

[1] Laboratory of Aquatic Ecology, Evolution & Conservation, Department of Biology, KU Leuven, Leuven, Belgium

[2] Laboratory of Reproductive Genomics, KU Leuven, ON I Herestraat 49, B-3000 Leuven, Belgium

[3] Department of Gynaecology and Obstetrics, University Hospitals Leuven, Leuven, Belgium

[4] Division of Gynaecological Oncology, Leuven Cancer Institute, KU Leuven, Leuven, Belgium

Modified from *Frontiers in Microbiology* 2019, 10:743

2.1 Abstract

Urbanization is transforming and fragmenting natural environments worldwide, driving changes in biological communities through alterations in local environmental conditions as well as by changing the capacity of species to reach specific habitats. While the majority of earlier studies have been performed on higher plants and animals, it is crucial to increase our insight on microbial responses to urbanization across different spatial scales. Here, using a metacommunity approach, we evaluated the effects of urbanization on bacterioplankton communities in 50 shallow ponds in Belgium (Flanders region), one of the most urbanized areas in Northwest Europe. We estimated the relative importance of local environmental factors (35 abiotic and biotic variables), regional spatial factors and urbanization (built-up area) quantified at two spatial scales (200 x 200 m and 3 x 3 km). We show that urbanization at local or regional scales did not lead to strong changes in community composition and taxon diversity of bacterioplankton. Urbanization at regional scale (3 x 3 km) explained only 2% of community composition variation while at local scale (200 x 200 m), no effect was detected. Local environmental factors explained 13% (OTUs with relative abundance ≥0.1%) to 24% (12 dominant OTUs - ≥1%) of community variation. Six local environmental variables significantly explained variation in bacterioplankton community composition: pH, alkalinity, conductivity, total phosphorus, abundance of *Daphnia* and concentration of copper (Cu), of which pH was partly mediated by urbanization. Our results indicate that environmental rather than spatial factors accounted for the variation in bacterioplankton community structure, suggesting that species sorting is the main process explaining bacterioplankton community assembly. Apparently, urbanization does not have a direct and strong effect on bacterioplankton metacommunity structure, probably due to the capacity of these organisms to adapt towards and colonize habitats with different environmental conditions and due to their fast adaptation and metabolic versatility. Thus, bacterioplankton communities inhabiting shallow ponds may be less affected by environmental conditions resulting from urbanization as compared to the impacts previously described for other taxa.

Keywords: urbanization, freshwater bacterioplankton, metacommunity, spatial scales, shallow ponds.

2.2 Introduction

The extent of anthropogenic transformation of biotic and abiotic conditions on Earth is so profound that we have now entered a new geological epoch, aptly dubbed the Anthropocene (Lewis & Maslin, 2015; Waters et al., 2016). With the dramatic expansion of urban land area and the global urban human population over the last few decades, urbanization has become one of the most extreme forms of land-use alterations triggered by humans (Grimm et al., 2008). Urban areas fundamentally differ from natural areas in terms of atmospheric chemistry, geochemistry, climate, hydrology and vegetation cover and thus form unique ecosystems (Alberti et al., 2003; Kaye et al., 2006). Consequently, there is a growing need to better understand the impact of urbanization on biodiversity and ecosystem functioning.

The relationship between urbanization and biodiversity is multifaceted. Urbanization can affect species richness, evenness and composition by means of several direct and indirect mechanisms. Habitat loss and fragmentation caused by urbanization might lead to reduced species richness and evenness thereby causing biotic homogenization (McKinney, 2006). An increase in impervious surface may also result in reduced species richness and changes in species composition through physical and chemical changes due to alterations of the hydrologic regime and polluted run-off (Morse et al., 2003; Chadwick et al., 2006). The introduction of exotic species by humans in cities (e.g. for cultivation purposes and by accidental importation) may cause replacement of native species (McKinney, 2008). In addition to changes in species composition, a recent meta-analysis has shown that urbanization increases the rates of phenotypic change across different taxa (Alberti et al., 2017).

However, depending on the taxonomic group, the level of urbanization and the spatial scale investigated, contrasting patterns in responses of community composition and trait change to urbanization have been reported (McKinney, 2008). The majority of earlier studies have been performed on higher plants and animals, with only a limited number of studies so far focusing on microbial responses to urbanization in soil (Xu et al., 2014; Reese et al., 2016), artificial surfaces (Afshinnekoo et al., 2015), streams (Teittinen et al., 2015) and lakes (Fisher et al., 2015; Newton & McLellan, 2015). Considering the tremendous diversity of bacteria (Hug et al., 2016) and their key role in major ecosystem services (Falkowski et al., 2008; Van Der Heijden et al., 2008), it is crucial to increase our insight on microbial responses to urbanization.

So far, studies on microbial diversity in the context of urbanization have primarily focused on describing patterns of species richness and composition in specific cities and towns (Xu et al., 2014; Afshinnekoo et al., 2015; Reese et al., 2016). In order to develop a cohesive theory on how microbes respond to urbanization, a better understanding of ecological processes underlying the observed responses is needed.

High dispersal rates, short generation times and rapid adaptation to environmental stressors are some of the important traits of microbes that set them apart from the majority of macro-organisms (Andersson et al., 2014). Species responses to urbanization have been shown to depend on dispersal rates and degree of habitat specialization that might in turn influence metacommunity dynamics, with implications at multiple spatial scales (Concepción et al., 2015). Urban ecological footprints can be observed at scales ranging from local (altered abiotic and biotic conditions) to regional (altered habitat heterogeneity and connectivity) and therefore biodiversity responses to urbanization should best be considered within a hierarchical metacommunity framework (Johnson et al., 2013). The metacommunity framework summarizes the relative contribution of local and regional processes in regulating the composition of ecological communities (Leibold et al., 2004). Most studies have reported species sorting as the predominant mechanism shaping microbial communities (Van Der Gucht et al., 2007; Souffreau et al., 2015; Muscarella et al., 2016). Species sorting emphasizes the role of local abiotic and biotic conditions in determining community composition and requires dispersal rates that are sufficiently high to allow for species to reach the habitats with their preferred niches but not so high to cause homogenization via mass effects (Leibold et al., 2004). Species sorting likely predominates in bacterial communities because of a combination of high dispersal rates and fast local population growth rates (Van Der Gucht et al., 2007). Some studies, however, have provided evidence for neutral dynamics, highlighting the role of regional factors in regulating community assembly in microbes (Drakare & Liess, 2010; Langenheder & Székely, 2011; Langenheder et al., 2012). Others observed spatially limited distribution patterns in microbial species, ranging from centimeters to hundreds of kilometers (Martiny et al., 2006) and being the result of historical processes (e.g. drift, priority effects and past environmental conditions) (Fukami et al., 2010; Andersson et al., 2014). In addition, a synergistic interaction of species sorting and neutral processes has also been detected (Langenheder & Székely, 2011; Liu et al., 2015).

Wetlands, ponds and lakes provide important ecosystem services to society and these services are to a large extent mediated by microbial communities (e.g., water purification, degradation of xenobiotics, and nutrient recycling) (Kremen, 2005). Ponds are ideal model systems for the study of ecology, evolution and conservation in a metacommunity context because they often show contrasting ecological conditions and have clearly delimited boundaries in the landscape (De Meester et al., 2005). Along an urbanization gradient, buildings, roads and other urban features might influence the dispersal of freshwater bacterioplankton. For example, buildings and other urban structures could block or difficult the microbial dispersal by wind. Besides, the effects of urbanization could have an indirect effect on microbial dispersal through the impact on vector species (e.g., water birds). Nevertheless, urbanization could also have a positive effect on microbial dispersal through facilitation of water run-off across different water bodies over impermeable surfaces such as roads and streets. In this context, a better understanding on how microbial communities in ponds and lakes are responding to the effects of urbanization at a metacommunity level is crucial.

In this study, we investigated the structure and composition of bacterioplankton communities in freshwater ponds located along well-characterized urbanization gradients in Belgium in order to get insight into the ecological processes acting on bacterioplankton in an urbanization context. We employed a spatially hierarchical sampling site selection to capture the combined as well as independent influence of environmental and spatial processes on bacterioplankton metacommunity structure across strong gradients of urbanization at two different spatial scales. Using freshwater ponds as study systems, we tested the follow hypotheses: (1) urbanization would lead to more homogeneous bacterioplankton community composition, with relatively minor effects on diversity; (2) in urban landscapes, environmental factors would be relatively stronger than spatial factors in regulating bacterioplankton community assembly. This implies that fragmentation effects captured by an increase in urbanization along the gradient (measured using the hierarchical design) are less important compared to changes in local conditions (abiotic and biotic factors) mediated by urbanization. These hypotheses are based on the assumption that urbanization has a stronger effect on bacterioplankton community composition rather than on diversity because some species might thrive under the urban conditions replacing other less adapted species, leading to species turnover rather than a loss in diversity. Moreover, it is presumed that changes in

environmental conditions related to urbanization at the local scale are more important to determine bacterioplankton community assembly compared to effects of urbanization at the regional scale, especially for microorganisms.

2.3 Materials and methods

Hierarchical sampling design and bacterioplankton sampling

Sampling was carried out within a polygon broadly defined by the cities of Gent, Antwerp, Leuven and Brussels (approx. 5000 km^2; Belgium) during summer 2013. A GIS-based (built-up area, i.e. percentage of area covered by buildings - data from *Grootschalig Referentie Bestand Vlaanderen* 2013; scale: 1/250-1/5000, Agency of Geographical Information Flanders [AGIV]) site selection was performed to select 27 plots of 3 km x 3 km, categorized based on the percentage of built-up area as high (> 15%; 9 plots), medium (5-10%; 9 plots) and low (< 3%; 9 plots) urbanized plots (a detailed map of the sampling design is available in Figure S1). Within each of these 27 plots, three subplots were selected representing the same three levels of urbanization (high, medium and low) at the scale of 200 x 200 m, resulting in 81 subplots. In this way, it was possible to capture the influence of three levels of urbanization (low, medium and high) at two different spatial scales. As built-up area is the area occupied by buildings, excluding streets, parking lots, etc., a > 15% built-up area represents a high level of urbanization. In each of the 81 subplots, a shallow pond (area < 1 ha, depth ≤ 3 m, fishless) was selected for sampling. After leaving out samples with insufficient coverage of sequences, 50 shallow ponds from the original design (81 shallow ponds) were analysed in this study. The resulting dataset is to some degree unbalanced towards the different urbanization categories, but the imbalance is moderate (see Table S1). Unfortunately, we do not have data on the management and function of the sample ponds used in this study.

The field survey was performed between end of May and beginning of July 2013, and each pond was sampled once. To sample water for characterization of bacterioplankton (defined as bacterial communities that drift along the water column) and abiotic variables, we collected a depth integrated water sample (in total 15 L) with a pre-rinsed tube sampler (length 1.85 m; diameter 75 mm) in the central area of each pond. We pre-filtered pond water using a 2 mm mesh size sieve to remove larger materials. Samples were kept at 4 °C in 2 L dark

jars until they were further processed at the laboratory on the same day of sampling. For bacterioplankton characterization, we filtered up to 100 mL of pond water through a MCE filter with 0.22 µm pore size (Milipore® - U.S.A.) and the samples were stored at -80 °C until DNA extraction.

Measurement of abiotic and biotic variables

To characterize local environmental biotic and abiotic factors, we utilized standardized protocols following Declerck *et al.* (2006) and De Bie *et al.* (2012), summarized as follows. In the field, pH, conductivity, water temperature and dissolved oxygen were measured using a Hach® multimeter (U.S.A.). For measuring transparency, we used a Snell tube, an adaptation of the classic Secchi disc method. In brief, we added water to a dark grey plastic tube (length 60 cm; diameter 8 cm) and assessed the depth at which a small 5 cm cross-section Secchi disk could be seen in that tube. This allows to measure transparency even in shallow systems where the Secchi depth would exceed water depth. Shaded pond area was visually estimated, and sludge depth, maximum depth and bank angle were measured using a graduated scale. Surface area of the ponds was estimated in Google Earth (version 7.1.8 - U.S.A.). A table with summarized pond morphometric data is available in Table S2.

We measured total phosphorus concentration of the water samples with the ascorbic acid method after perchlorate digestion (Murphy & Riley, 1962) and total nitrogen concentration using the Kjeldahl method (United States Environmental Protection Agency (USEPA), 1983). The concentration of available phosphorus and nitrogen (dissolved nutrients) was determined after filtration of 50 mL of water on a GF/F filter with a Technicon autoanalyser III (United Kingdom). Dissolved metals (Cu, Ni, Zn, Al, Cr, Mg, Fe, Mo, As, Ag, Cd and Pb) were quantified using a ICP-MS (Inductively Coupled Plasma Mass Spectrometry) (Agilent 7700x ICP-MS) (U.S.A.) after filtration over a 0.22 µm pore size sterile filter. Alkalinity and the concentration of sulfates were measured in the laboratory following standardized protocols using Hach® Water Analysis (Hach, 1992) (U.S.A.). Dissolved organic carbon was estimated as the NPOC (Non-purgeable Organic Carbon) by catalytic reaction and infrared detection of the carbon dioxide formed after filtration of 15 mL of water on a 0.45 µm pore size filter. For the measurement of suspended matter, we filtered water samples (up to 1 L) over a GF/F filter that was previously dried at 105° C for six hours and weighted. After filtration of the water samples, the filters were dried again following the same procedure and the

difference of weight between the pre-weighted and the dried filter was recorded as suspended mater content.

We also incorporated biotic variables in our analyses: concentration of chlorophyll-a (estimated with a calibrated Aquafluor® Turner Designs fluorometer – U.S.A.), zooplankton abundance, *Daphnia* abundance, biomass of *D. magna*, and *Ceriodaphnia* abundance. The inclusion of zooplankton as explanatory variables is based on previous knowledge on the importance of certain genera such as *Daphnia* in structuring bacterioplankton communities (Langenheder & Jürgens, 2001; Verreydt et al., 2012; Berga et al., 2015). We sampled zooplankton using a tube sampler (length 1.85 m; diameter 75 mm), collecting in total 60 L of depth-integrated water sample at the pelagic and littoral zones from each pond. We filtered 20 to 40 L of this volume over a 64 µm mesh size sieve. Samples were fixed with formalin (7%) and stored in a 60 ml vials for species identification in the laboratory. Zooplankton identification was conducted using a stereomicroscope and a minimum of 300 individuals were identified per sample (Gianuca et al., 2018). Zooplankton abundances were calculated as number of individuals per litter sample. Zooplankton, bacterioplankton and environmental sampling were performed on the same day for a specific pond.

DNA extraction, PCR of 16S rRNA genes and sequencing

DNA was extracted using the UltraClean™ Soil DNA isolation kit (MoBio Laboratories Inc., Carlsbad, CA, U.S.A.) and amplified by PCR using the 16S rRNA gene universal eubacterial primers E338F (5'-ACTCCTACGGGAGGCAGCAGT-3') and E797R (5'-GGGTATCTAATCCTG-3') (amplicon length of 459-bp covering the variable V3 region of the 16S rRNA gene). The PCR mixtures were prepared with 2.5 µl 10 x PCR buffer (Eurogentec, Seraing, Belgium), 1 µl $MgCl_2$ (50 mM; Eurogentec), 2.5 µl dNTPs (2 mM; Thermo Scientific, MA), 1.0 µl of each primer (20 pmol μl^{-1} each; Eurogentec), 0.2 µl Silverstar Taq DNA polymerase (5.0 U μl^{-1}; Eurogentec) and 1 µl standardized template DNA (10.74 ng μl^{-1}) in a total reaction volume of 25 µL. The PCR was performed with the follow temperature/time conditions: initial denaturation at 94 °C for 2 min, 30 cycles of 94 °C for 1 min, 55 °C for 1 min and 72 °C for 1 min, and a final extension for 5 min (72 °C). PCR products were evaluated on gel electrophoresis, the bands then were excised and purification was performed using a Qiaquick Gel extraction kit (Qiagen, Venlo, Netherlands). For the quantification of the PCR products, we used the PicoGreen dsDNA Assay

Kit (Life Technologies - U.S.A.). Sequencing was performed at the Genomics Core of KU Leuven (Belgium) using a Roche 454 GS FLX+ (Switzerland).

Processing 454 pyrosequencing data

We processed the 16S rRNA gene pyrosequences using the software mothur (version 1.30.2 - U.S.A.). The sequences contained in the flowgrams (SFF file) were sorted into groups according to the barcode sequences. The standard operating procedure (http://www.mothur.org /wiki/Schloss_SOP) was followed for the analysis (Schloss et al., 2011). A total of 89015 good-quality sequences were recovered from the 50 samples that were used for downstream analysis. A 97% identity in rRNA gene sequence was used as the operational taxonomic unit (OTU) definition. The sequences were assigned using as reference the SILVA database. A total of 1133 OTUs (no global singletons) were identified from the total set of samples. Samples were rarefied to 688 sequences per sample to compare relative differences between samples. Raw sequences from all samples were submitted to the European Nucleotide Archive under the accession numbers ERR1871981-ERR1872018.

Statistical analysis

We conducted all analyses in R (version 3.2.2, 2014) using the package vegan (Oksanen, 2015). Except for richness, all analyses were performed on relative abundance data. Diversity and exclusiveness analyses were performed using all 1133 OTUs (no global singletons). Since the sequencing depth was relatively low in our study, we assume that we mainly captured the most abundant taxa in our samples. We therefore analysed these data making two subsets, with different cut-off levels: ≥0.1% (abundant taxa - 100 OTUs) and ≥1% (dominant taxa - 12 OTUs) of the rarefied total dataset. Additionally, this could be justified because microbial communities are usually composed of a small group of abundant organisms and a vast number of rare taxa (Pedrós-Alió, 2012; Baltar et al., 2015; Liu et al., 2015). To estimate the impact of urbanization, all the below-mentioned analyses were performed for the two datasets. In order to visualize compositional changes among the different urbanization categories at plot and subplot levels, we used a non-metric multidimensional scaling (NMDS) approach with Bray-Curtis index as measure of dissimilarity. To verify the changes in composition due to urbanization and associated environmental variables, we performed PERMANOVA (Anderson, 2001) using the *adonis* function (number of permutations = 9999, stratum = plot identifier –

this parameter was included to account for the spatial dependency of subplots within the same plot) of the vegan package. To evaluate whether variability in composition within samples changed among urbanization categories, we performed PERMDISP (Anderson, 2006) using the function *betadisper* followed by a permutation test of the F values using the function *permutest*. PERMANOVA and *permutest* analyses were performed using 9999 permutations. The effect of urbanization level at both the local (subplot) and regional (plot) scale and their interaction on the (i) diversity, (ii) abundances of each of the ten most abundant classes, (iii) abundance of each of the 12 dominant OTUs, and (iv) six selected environmental variables was tested univariately by means of generalized linear mixed models (GLMMs) performed with the function *lme*. Levels of urbanization (low, medium, high) at the subplot and plot scale and their interaction were set as fixed factors. To account for the spatial dependency of subplots within the same plot, a plot identifier was incorporated as a random factor, nested within plot. The tested environmental variables were selected by forward selection for the variation partitioning analysis described below.

The relative contribution of spatial (S) and environmental (E) factors, urbanization level at plot level (U_{plot}) and urbanization at subplot level ($U_{subplot}$) in determining local species composition was determined using a variation partitioning approach (Peres-Neto et al., 2006). To identify which variables should be incorporated in the variation partitioning for the environmental and spatial models, we used forward selection (Blanchet et al., 2008). Only significant variables selected by the forward selection (permutation test = 9999, $P < 0.05$) were included in the environmental (E) and the spatial models (S) for the variation partitioning analysis. For the variation partitioning and the forward selection analyses, we used the function *varpart* and *forward.sel*. To assess the pure components (i.e. corrected for the other components, e.g. E|(S + U) is the pure environmental component without spatial and urbanization effects), a permutation test (1000 permutations) was performed using the functions *rda* and *anova.cca*. We also used the *by=terms* command to test which specific variables contributed to the pure environmental model. All environmental variables were first tested for possible collinearity based on Pearson correlation values. When the correlation between two variables was higher than 0.7, only one of the variables was included in the downstream analyses. The selection of variables was also based on existing knowledge on the importance of the different variables in an urbanization context. All environmental variables were log-transformed (except pH) and standardized (z-scores). The spatial component was

generated based on the geographical coordinates of each site. To incorporate the non-linear spatial components, we used Moran's eigenvector map (MEM) (Dray et al., 2006). Latitude and longitude were included as variables to test separately for the linear spatial component. Urbanization levels (three categories) were given as dummy variables. For the community data, we used Hellinger transformation to reduce the impact of zero values in our dataset (Legendre & Gallagher, 2001).

2.4 Results

Urbanization-driven community composition and diversity

Despite the low number of sequences per sample (688 sequences), almost the totality of samples reached saturation for the abundant OTUs (relative abundance ≥0.1% of the total dataset) (Figure S2). This indicates that the abundant fraction of the communities could be accurately characterized using this number of sequences. Besides, performing all the analyses with a coverage of 1000 sequences per sample (38 samples) gave similar results (Table S3 to Table S5).

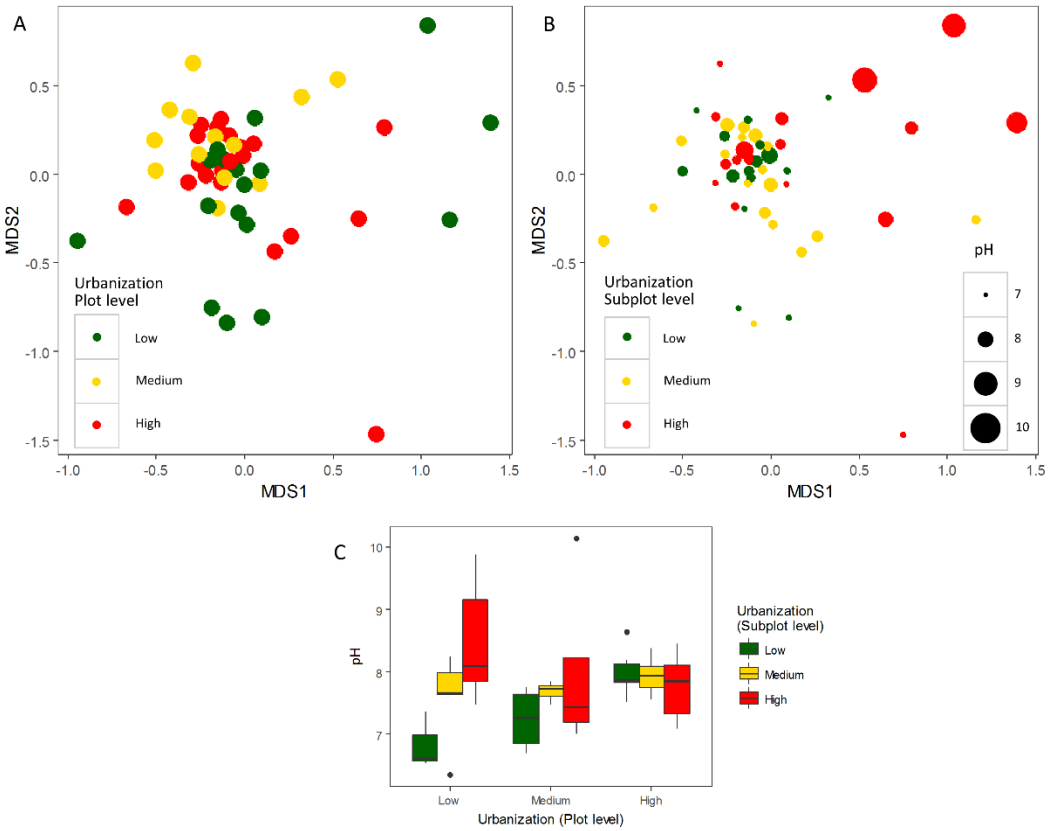

Figure 1. Non-metric multidimensional scaling (NMDS) plot of bacterioplankton communities (abundant OTUs - relative abundance ≥0.1%) coloured by urbanization categories (low, medium and high) at plot level (A) and subplot level (B). Urbanization categories are represented by colours and pH values represented by the size of each dot for the subplot level. (C) Boxplots of pH values across urbanization categories (low, medium and high) at subplot and plot levels.

An effect of urbanization category (low, medium and high) on bacterioplankton community composition was detected at plot level by PERMANOVA (R^2 = 0.06, *P* = 0.01) [visualized by NMDS (STRESS = 0.18) in Figure 1A], whereas at subplot level, no effect of urbanization was detected (R^2 = 0.04, *P* = 0.27) [visualized by NMDS (STRESS = 0.18) in Figure 1B]. No interaction between subplot and plot levels of urbanization on bacterioplankton community composition was detected (R^2 = 0.84, *P* = 0.30) (Table S6).

We did not detect any significant effect of urbanization at subplot or plot levels on bacterioplankton diversity (number of OTUs, Shannon diversity, Chao1 and ACE estimators; Table S7 and Figure S3). To estimate whether more subtle changes could be detected at more specific levels, we tested the effect of urbanization on relative abundances of the most dominant OTUs, and also separately on phylum and class levels. A total of 17 phyla were detected in the studied ponds for the abundant fraction (relative abundance ≥0.1% of the total dataset). Proteobacteria (44.44% of total abundance), Actinobacteria (21.32%), Bacteroidetes (16.63%) and Firmicutes (1.95%) were the most abundant phyla. No significant differences in the relative abundances across urbanization categories were detected for the ten most abundant classes found in our study at both scales studied (local and regional) (Figure 2C – four most abundant classes depicted) (Table S8 and Figure S4). Among the OTUs restricted to highly urbanized subplots that were identified to genus level, *Algoriphagus*, *Pedobacter* and *Flavobacterium* were the most abundant (Table 1). Among the 12 dominant OTUs (>1% of relative abundance), which accounted for 57.92% of the total community composition, two OTUs showed significant differences in relative abundance across urbanization categories (Figure S5). An OTU from the class Actinobacteria was more abundant in high urbanization subplots than in medium and low urbanizations subplots and was more abundant in medium urbanization plots. An OTU from the genus *Methylobacter* (Table S9) was less abundant in medium than in low and high urbanization subplots (Figure S5).

Table 1. Bacterioplankton taxa that are restricted (> 90% of occurrence) to ponds in highly urbanized subplots (# reads = total number of reads in all ponds). Taxa are ranked by percentage of exclusiveness (EXC) (ex: high – 100% means only occurrence in ponds situated in highly urbanized subplots).

phylum	class	order	family	genus	# reads	EXC (%)
Bacteroidetes	Sphingobacteria	Sphingobacteriales	Cyclobacteriaceae	*Algoriphagus*	38	100
Bacteroidetes	Sphingobacteria	Sphingobacteriales	Sphingobacteriaceae	*Pedobacter*	35	100
Proteobacteria	Epsilonproteobacteria	Campylobacterales	Helicobacteraceae	*Sulfurimonas*	28	100
Bacteroidetes	Flavobacteria	Flavobacteriales	Flavobacteriaceae	*Flavobacterium*	14	100
Actinobacteria	Actinobacteria	Actinomycetales	unclassified	unclassified	14	100
Bacteroidetes	unclassified	unclassified	unclassified	unclassified	13	100
Bacteroidetes	Flavobacteria	Flavobacteriales	Flavobacteriaceae	unclassified	12	100
Bacteroidetes	Sphingobacteria	Sphingobacteriales	Chitinophagaceae	unclassified	11	100
Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	*Sandarakinorhabdus*	10	100
Bacteroidetes	unclassified	unclassified	unclassified	unclassified	10	100
Bacteroidetes	Sphingobacteria	Sphingobacteriales	Chitinophagaceae	unclassified	10	100
Proteobacteria	Betaproteobacteria	Burkholderiales	Burkholderiaceae	*Limnobacter*	98	98.9
Proteobacteria	unclassified	unclassified	unclassified	unclassified	56	98.2
Proteobacteria	Epsilonproteobacteria	Campylobacterales	Helicobacteraceae	*Sulfurimonas*	82	96.3
unclassified	unclassified	unclassified	unclassified	unclassified	72	95.8
Fusobacteria	Fusobacteria	Fusobacteriales	Fusobacteriaceae	*Cetobacterium*	21	95.2
Proteobacteria	Alphaproteobacteria	Rhodospirillales	Acetobacteraceae	*Roseomonas*	39	94.9
Bacteroidetes	Sphingobacteria	Sphingobacteriales	Chitinophagaceae	*Sediminibacterium*	45	93.3
Proteobacteria	Alphaproteobacteria	Rhizobiales	unclassified	unclassified	39	92.3
Bacteroidetes	Flavobacteria	Flavobacteriales	unclassified	unclassified	13	92.3
Firmicutes	Bacilli	Bacillales	Bacillaceae_1	*Bacillus*	13	92.3
Bacteroidetes	unclassified	unclassified	unclassified	unclassified	21	90.5

In order to estimate other effects of urbanization on bacterioplankton structure, we estimated the number of exclusive OTUs along the gradient and pointed out the most common OTUs at highly urbanized subplots. At the subplot (local 200 x 200 m) level, there were a higher number of exclusive OTUs in high urbanization ponds (235) compared to low (108) and medium (159) urbanization ponds (Figure 2A). At the plot (regional 3 x 3 km) level, there were a higher number of exclusive OTUs in low urbanization ponds (231) compared to high (212) and medium (80) urbanization ponds (Figure 2B). For the OTUs exclusive to ponds in highly urbanized subplots, the phylum Bacteroidetes was represented by the families Cyclobacteriaceae, Sphingobacteriaceae, Flavobacteriaceae, Chitinophagaceae and Sphingomonadaceae (Table 1).

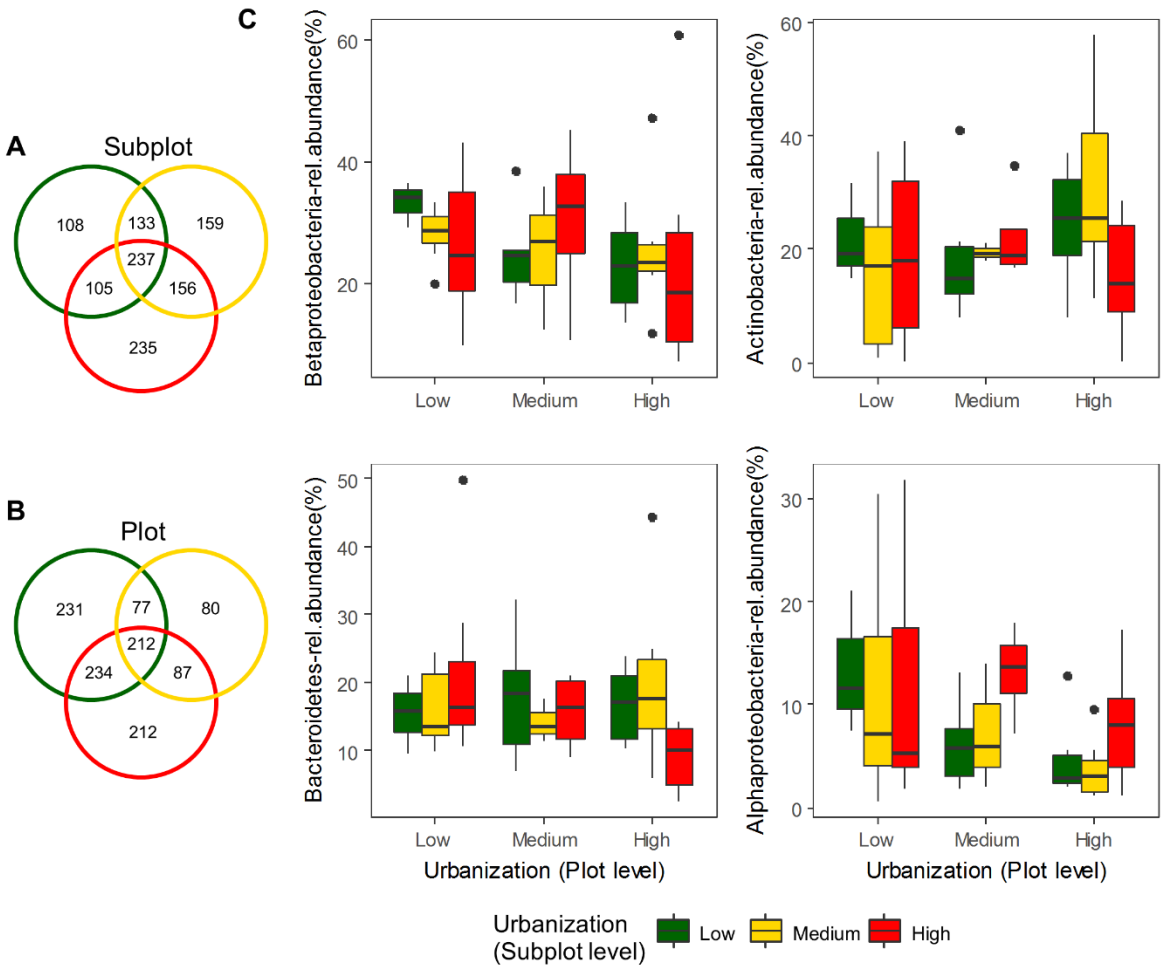

Figure 2. Venn diagrams showing exclusive and shared OTUs for the three urbanization categories (total dataset - 1133 OTUs) at (A) subplot level and (B) plot level. Boxplots of the four most abundant classes of bacteria found along the three urbanization categories (low, medium and high) at subplot (C) and plot levels (D).

Relative impact of urbanization at local and regional scale on community composition

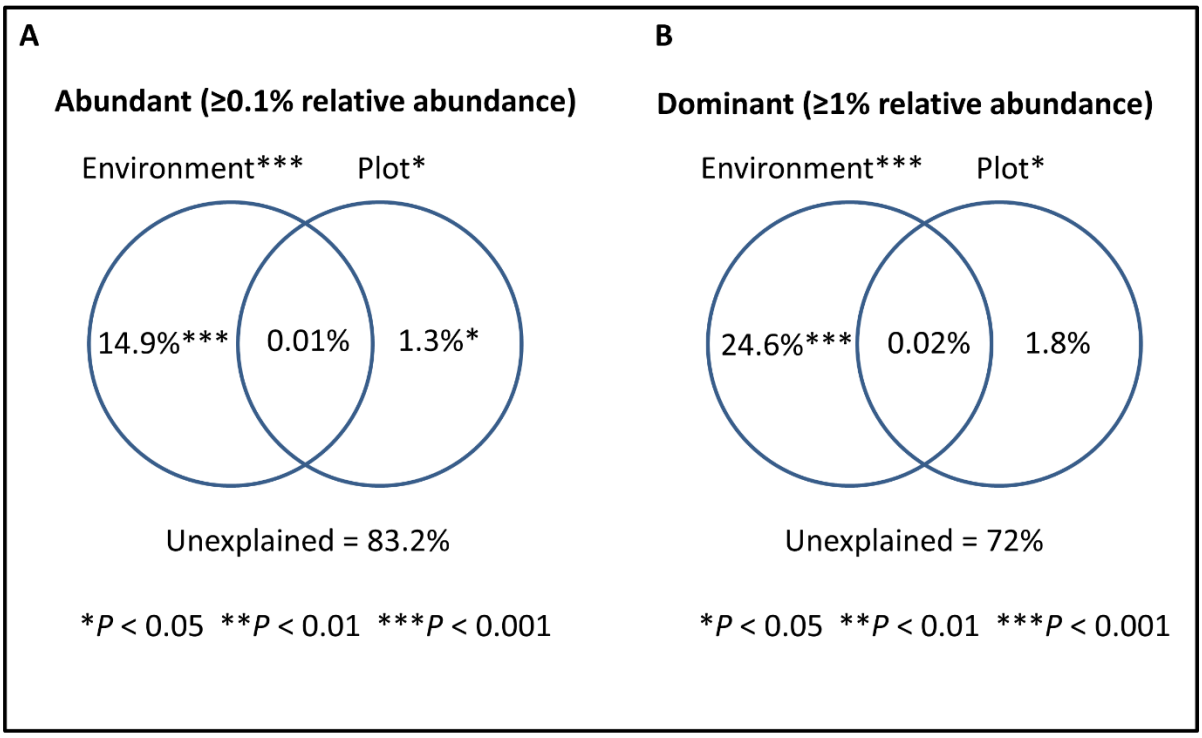

Figure 3. Venn diagrams showing the results of variation partitioning analyses. Percentage of variance explained by different models for the abundant fraction (100 OTUs) is shown in (A) and for the dominant OTUs of bacterioplankton community (12 OTUs) is shown in (B).

To test the relative contribution (adjusted R^2 values) of urbanization, environmental and spatial factors in shaping bacterioplankton communities, we performed a variation partitioning analysis. Environmental factors (variables selected by forward selection) and urbanization at plot level explained together 16.8% of the variation in bacterioplankton community composition of the abundant fraction (relative abundance ≥0.1%) (Figure 3A). In this model, pure environmental factors explained the highest amount of variation (14.9%, $P < 0.001$), while pure urbanization at plot level explained a small but significant part of the variation (1.3%, $P < 0.01$) (Table 2). For the 12 dominant OTUs (>1% relative abundance), environmental variables and urbanization at plot level explained together 28% of the variation (Figure 3B). Pure environment was the component explaining the largest amount of community variation for the dataset on dominant OTUs (24%, $P < 0.001$) (Table 2). Pure urbanization at plot level (1.8%) was not significant and 1.6% of the variation was shared between environmental factors and urbanization at plot level.

Table 2. Summary of the RDA analyses showing the relative importance of the different components of explained variation of community composition in the abundant (100 OTUs - relative abundance ≥0.1%) and in the dominant fraction (12 OTUs - relative abundance ≥1%). The effects of each component (E – environment, S – space, U – urbanization at subplot and plot levels) are not (marginal effect) or are corrected (conditional effect) for the other components. NS means that a certain model did not explain the variation in community composition significantly, or that the % explained variance was not significant.

Model	Components studied							
	Marginal effect				Conditional effect			
	E	S	U(subplot)	U(plot)	E	S	U(subplot)	U(plot)
Abundant (100 OTUs)	0.154***	NS	NS	0.019*	0.149 ***	NS	NS	0.013*
Dominant (12 OTUs)	0.261***	NS	NS	0.034*	0.246***	NS	NS	0.018^{NS}

E – Environment, S – Space, U – Urbanization (NS – not significant, ***$p < 0.001$, **$p < 0.01$, *$p < 0.05$)

Environmental factors correlated with community structure

Six local environmental variables were selected by forward selection to significantly explain variation in bacterioplankton community composition: conductivity, pH, alkalinity, total phosphorus, abundance of *Daphnia* and concentration of copper (Cu) (Table 3). Based on visual inspection of Figure 4A, bacterioplankton communities in ponds located in highly urban (at subplot level) areas were positively related with elevated pH values (Figure 4A). In addition, bacterioplankton communities in highly urbanized areas have a significantly greater variation among themselves compared to those in areas of medium and low urbanization (Figure 4B). When corrected for space and urbanization (at both scales), five variables were retained for the conditional environmental model (E|S+U) of the abundant fraction: pH, alkalinity, total phosphorus, abundance of *Daphnia* and concentration of copper (Cu) (Table 3). For the dataset considering only the 12 most abundant OTUs, seven local environmental variables were selected for the uncorrected environmental model: pH, abundance of *Daphnia*, suspended matter, total nitrogen, maximum depth, molybdenum (Mo) and copper concentration (Cu) (Table 3). When corrected for urbanization at plot level, the conditional environmental model still retained all variables (Table 3).

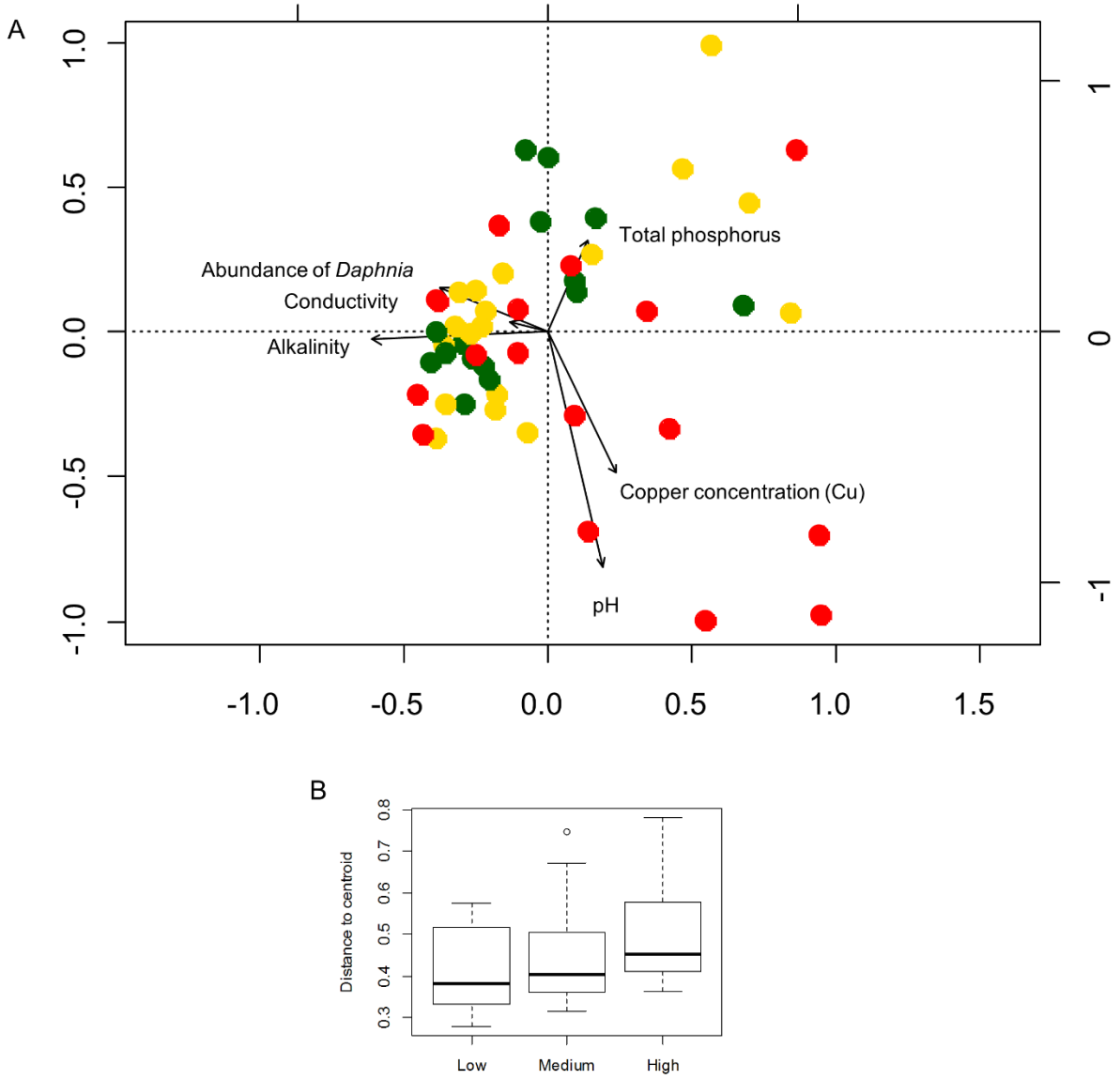

Figure 4. (A) Redundancy analysis of the bacterioplankton community composition for the abundant fraction (100 OTUs) and environmental variables selected by forward selection. Colors represent levels of urbanization around the ponds (low, medium, and high level) at subplot level. (B) Box-plot showing the variation (in terms of beta diversity) within bacterioplankton communities among the urbanization categories at subplot level. The average distance of a sample to the centroid of each urbanization category was used to calculate an F-statistic followed by a PERMUTEST.

Since pH was the variable explaining the highest variation in bacterioplankton communities (3.5% - abundant, 6.5% - dominant dataset), we further analyzed the effect of pH on bacterioplankton community composition along the studied urbanization gradient. Bacterioplankton community composition significantly varied along the pH gradient as confirmed by PERMANOVA (for categorical classification of pH < 7, 7-8, 8-9 and > 9, PERMANOVA, $P < 0.01$). Ponds with the highest pH levels, mostly in subplots characterized by

a high level of urbanization, harbored clearly distinct bacterioplankton communities (Figure 1B). Further, we found a significant difference among pH values across urbanization subplots, with higher pH values in high urbanization subplots than in low urbanizations subplots (P = 0.024) (Figure 1C) (Table S10). The other environmental variables selected by forward selection driving bacterioplankton community composition did not show significant differences across urbanization plots and subplots (Figure S6).

Table 3. Results of RDA analyses. Significant environmental variables identified by forward selection, explaining the variation (adjR2) in bacterioplankton community composition. Environmental variables effects were not (marginal effect) or were corrected (conditional effect) for the abundant (100 OTUs - relative abundance ≥0.1%) and in the dominant fraction (12 OTUs - relative abundance ≥1%).

Model	Significant environmental variables			
	E (marginal effect)	adjR^2	E|S+U (conditional effect)	adjR^2
Abundant (100 OTUs)	pH	0.035**	pH	0.027***
	alkalinity	0.036***	alkalinity	0.029***
	total phosphorus	0.024**	total phosphorus	0.016**
	abundance of *Daphnia*	0.024**	abundance of *Daphnia*	0.020**
	conductivity	0.019**	copper (Cu)	0.014*
	copper (Cu)	0.013*		
Dominant (12 OTUs)	pH	0.065***	pH	0.029***
	abundance of *Daphnia*	0.047**	suspended matter	0.012*
	suspended matter	0.044**	total nitrogen	0.010*
	molybdenum (Mo)	0.040**	maximum depth	0.013*
	copper (Cu)	0.028*	abundance of *Daphnia*	0.016**
	total nitrogen	0.016*	copper (Cu)	0.010*
	maximum depth	0.021*	molybdenum (Mo)	0.013*

E – Environment, E|S + U – Environment corrected for space and urbanization (plot and subplot)
(***p < 0.001, **p < 0.01, *p < 0.05)

2.5 Discussion

We evaluated how and to what extent bacterioplankton communities of shallow ponds are affected by urbanization in a metacommunity context. Based on 50 shallow ponds sampled in Belgium, urbanization (measured as percentage of built-up area) turned out to have only a minor direct effect on bacterioplankton community composition. At local subplot level (200 x 200 m), urbanization did not explain any significant portion of community variance, and at regional plot level (3 x 3 km), urbanization explained only a small significant fraction of variation (1.9% for the abundant fraction, 3.4% for the dominant fraction). In contrast, the measured environmental variables explained 15.4 - 26.1% of the variation in bacterioplankton when corrected for confounding effects of urbanization. As we predicted, urbanization had an indirect effect on bacterioplankton community, apparently mediated by environmental factors. Also, at the level of OTU diversity, no significant effects of urbanization were observed, and relative abundances of the most common bacterioplankton classes did not differ among the urbanization categories, although significant differences were observed for two dominant OTUs.

These results are in contrast to the relatively strong effects of urbanization reported for birds, mammals and plants, which include biotic homogenization (McKinney, 2006) and reduced biodiversity (Grimm et al., 2008). The bacterioplankton communities in our study do not follow these trends, as we observed a significantly higher variation in bacterioplankton community composition among urban compared to rural and semi-natural sites at the subplot level (Figure 4B) and no effect on diversity on any of the spatial urbanization scales (Figure S3). We see different potential reasons for this contrasting response to urbanization of bacterioplankton in shallow ponds. First, most ponds sampled in this study were man-made ponds and probably managed. Although we do not have management data of the sampled ponds, we know that management can influence pond variables such as nutrient state and macrophyte abundance (Declerck et al., 2006). Variation in pond management among ponds might have resulted in a similar gradient in pond types in both urban and rural areas, independent of urbanization category. We indeed observed a similar gradient in environmental conditions in both rural, semi-urban and urban ponds (Table S11). Second, ponds are inherently highly fragmented systems and can be seen as islands in a terrestrial matrix. Habitat fragmentation and reduced dispersal, associated with urbanization and impacting many terrestrial organisms (Alberti et al., 2017), might play a less important role in community assembly of bacterioplankton and small aquatic organisms than of larger

organisms. It has recently been shown across studied areas in Manhattan (New York City, U.S.A.) that microbial richness was not reduced by the effects of urbanization, while it had a pronounced effect on macro-organisms (Reese et al., 2016).

The small proportion of variation in community composition explained by urbanization at plot level was mostly shared with environment variables (1.9 - 3.4%). One environmental variable (pH) was correlated with the level of urbanization and did influence bacterioplankton community composition significantly (Figure 1C). pH is known to be an important driver for microbial communities related to species turnover in extreme conditions (Ren et al., 2015). It also plays a role in continuous gradients such as presented here and also observed in other studies (Mukherjee et al., 2014; Xu et al., 2014; Yan et al., 2016). In our dataset the pH ranged from 6.35 to 10.14, playing a significant role in structuring the bacterioplankton communities studied here. Urbanization is known to mediate pH values due to an increase of impervious surfaces in more urbanized areas (Conway, 2007). This increase could be related to the leaching of bicarbonates of concrete structures used in impervious surfaces in urban areas (Davies et al., 2010).

Also, *Daphnia* abundance contributed significantly to explain variation in bacterioplankton community composition. Experimental studies have previously found that the abundance of *Daphnia* can have an impact on bacterioplankton communities (Langenheder & Jürgens, 2001; Verreydt et al., 2012; Berga et al., 2015). Large-bodied water fleas of the genus Daphnia are pivotal species in the food web of ponds and lakes, as they can exert a strong top-down impact on microbial communities (Jürgens & Jeppesen, 2000; Degans & De Meester, 2002; Zöllner et al., 2003). Moreover, a symbiotic relationship between *Daphnia* species and some important genera of bacterioplankton has been reported (Zöllner et al., 2003; Qi et al., 2009; Callens et al., 2016). Some bacteria, such as of the genus *Limnohabitans* (one of the most abundant taxa found in our dataset), are an important component of freshwater bacterioplankton communities and have been found to be also part of the microbiota of *Daphnia* species (Callens et al., 2016). A parallel analysis of metacommunity structure of zooplankton in the same systems revealed a small but significant signature of urbanization on the abundance of Daphniidae (Gianuca et al., 2018). The selection of *Daphnia* abundance as a factor to explain variation in bacterioplankton community composition (even when corrected for urbanization) reinforces the importance to include biotic variables to better understand what drives bacterioplankton communities.

While other general pond characteristics such as conductivity, nutrients concentration and abundance of *Daphnia* highly structured the bacterioplankton communities in the present dataset independently of urbanization, there are some small effects correlated to urbanization level. These effects on bacterioplankton community composition are less strong than we would have expected based on the studies in terrestrial and other aquatic organisms that were conducted within the same hierarchical design and field survey as the present study. A study on zooplankton using the same ponds also found that the effect of urbanization on community composition was significant but small (4%; Gianuca *et al.*, 2018). In terrestrial carabid beetles, using slightly different subplots, changes in species composition were significantly related to urbanization at both spatial scales (Piano et al., 2017). Finally, a study analysing community data from this same design across ten macro-organisms found a strong effect of urbanization in shifts of body size (Merckx et al., 2018). It has recently been shown that the effects of urbanization on microbial organisms might be different compared to macro-organisms, where urbanization leads to reduced diversity and homogenization (Reese et al., 2016). For aquatic organisms such as micro-organisms that are not strongly dispersal limited, the ponds and lakes in urban areas can represent a broad range of environmental conditions due to the interplay of different abiotic and biotic factors and the release from interference from larger organisms which could allow more or different species to coexist. The level of exclusiveness for some taxa found in urban ponds further reinforces this notion (Figure 2A and 2B).

Our variation partitioning analyses further indicate that environmental rather than spatial factors accounted for the variation in bacterioplankton community structure. This suggests that species sorting was the main process explaining bacterioplankton community assembly, reinforcing previous observations that abiotic and biotic variables have a stronger effect on bacterioplankton community variation than regional factors across different spatial scales (Van Der Gucht et al., 2007; Jones & McMahon, 2009; Logue & Lindström, 2010; Souffreau et al., 2015). Urbanization seems to have an indirect effect mediating the relative importance of species sorting through water chemistry at subplot level. However, at plot level we detect a small, but direct effect on composition that could be related to an effect of urbanization on the bacterioplankton regional pool of species. These different effects of urbanization at subplot and plot level could suggest that urbanization act in distinct ways depending of the spatial scale. The relative importance of species sorting driving

bacterioplankton communities has been related to the degree of environmental variation among sites (Östman et al., 2010), with higher environmental heterogeneity leading to an increase in the relative importance of species sorting. The selection of 6-7 different environmental variables in our environmental model show that along the studied urbanization gradient, multiple environmental variables were involved in explaining changes in community composition of bacterioplankton.

The relatively higher proportion of explained variation, when only considering the dominant taxa (relative abundances ≥1%) compared to the abundant taxa (relative abundances ≥0.1) is likely due to statistical constraints, since it is known that microbial communities have an elevated number of species compared to macro-organisms and consequently the inclusion of rare species, introduces noise to the patterns (Siqueira et al., 2012; Székely & Langenheder, 2014). It has also been shown before that the dominant patterns in the structure of microbial communities are indeed governed by a relatively small group of taxa (Székely & Langenheder, 2014). The selection of a similar set of environmental conditions determining bacterioplankton structure for both fractions studied (Table 3) reinforces this idea.

Based on our results, bacterioplankton communities in shallow ponds along an urbanization gradient were structured by processes (i.e., species sorting) predominantly acting at the level of the individual pond. This result is similar to the one previously found with respect to agricultural land use intensity structuring different aquatic organisms in the same sampling area (De Bie et al., 2012). Environmental conditions may be more important than neutral processes for these aquatic microorganisms. pH was shown to be an important environmental factor and was mediated by urbanization. The greater dispersal capacity of microbes compared to macro-organisms could make them less susceptible to impacts of landscape fragmentation due to urbanization. Instead, changes in local conditions mediated by urbanization have a greater impact on bacterioplankton community composition. Species sorting is more likely the most important process explaining community variation of bacterioplankton along urbanization gradients. Analysis of the functional gene composition of bacterioplankton communities through metagenomics would give us more in-depth information on the impact of urbanization on bacteria beyond taxonomy.

Acknowledgements

We thank Matthias Vanhamel, Edwin van den Berg, Andros T. Gianuca and Carla Denis for their contributions to the field work and laboratory analyses. We also thank the different SPEEDY partners for their help with the landscape analyses and the contributions to the general development of the SPEEDY project. This work is part of the SPEEDY ("SPatial and environmental determinants of Eco-Evolutionary DYnamics: anthropogenic environments as a model") project funded by the Belgian Science Policy Office – Interuniversity Attraction Poles program [P7/04] and was further financially supported by KU Leuven Research Council excellence financing [PF/2010/07]. FTTH was supported by the Science without Borders program [process number: 45968/2012-1] of Conselho Nacional de Desenvolvimento Científico e Tecnológico - Brazil. SM was supported by a Marie Skłodowska Curie Individual Fellowship (MSCA-IF-EF-ST) by the European Commission [Project ID: 655159]. JE received a scholarship from the Institute for the promotion of Innovation by Science and Technology in Flanders [IWT - project number 121625]. KIB was supported by Fund for Scientific Research - Flanders [FWO Vlaanderen, application number 11N5116N].

2.6 Supplementary material for Chapter II

Table S1. Summary of the number of plots, subplot and combination of both levels of urbanization for the 50 sites sampled.

	Urbanization levels		
	Low	Medium	High
Plot	17	13	20
Subplot	15	18	17
Combination	Low (Plot)	Medium (Plot)	High (Plot)
Low (Subplot)	3	7	7
Medium (Subplot)	6	3	4
High (Subplot)	6	8	6

Table S2. Morphometric data for the 50 sites sampled.

	Morphometric variables			
Pond codes	Depth (m) / Transparency (m)	Surface area (m^2)	Average bank angle (cm)	Average depth at edge (cm)
PL1-YEL	1.5 / 0.5	1306	70	60
PL10-GRE	0.75 / 0.25	12	25	5
PL10-RED	1.3 / 0.30	31	40	15
PL10-YEL	1.2 / 0.16	589	90	40
PL11-GRE	1.42 / 0.43	294	15	10
PL11-RED	1.1 / 0.35	1225	65	15
PL11-YEL	1.4 / 0.29	628	45	15
PL12-GRE	0.7 / 0.17	75	40	10
PL12-RED	2 / 0.22	1200	35	10
PL13-GRE	1.8 / 0.18	1922	10	10
PL13-RED	0.7 / 0.05	117	5	10
PL13-YEL	0.6 / 0.5	314	30	10
PL14-RED	1.2 / 0.05	54	55	25
PL14-YEL	2 / 0.22	109	60	15
PL15-GRE	1 / 0.14	365	10	10
PL15-RED	1 / 0.27	1963	40	10
PL15-YEL	1.8 / 0.21	292	80	50
PL16-GRE	0.6 / 0.30	87	40	8.5
PL16-RED	0.3 / 0.21	176	90	5
PL18-YEL	0.65 / 0.24	471	35	7

PL2-GRE	1.5 / 0.22	82	35	30
PL2-RED	1.5 / 0.19	1100	45	20
PL2-YEL	1.5 / 0.43	375	80	20
PL23-GRE	1 / 0.12	58	75	12
PL23-RED	0.75 / 0.29	56	90	70
PL23-YEL	2 / 0.18	6400	30	10
PL24-RED	1 / 0.26	151	50	20
PL24-YEL	1.5 / 0.22	31	50	25
PL25-RED	0.7 / 0.11	56	60	30
PL25-YEL	0.7 / 0.28	95	40	14
PL26-GRE	0.85 / 0.22	39	20	8
PL26-RED	1.75 / 0.20	3675	35	20
PL26-YEL	1 / 0.17	628	15	13
PL27-GRE	0.7 / 0.11	282	50	7
PL27-RED	1 / 0.13	0.047	5	0
PL27-YEL	1 / 0.13	62	45	10
PL3-GRE	0.8 / 0.23	37	15	15
PL3-RED	1 / 0.25	157	15	15
PL3-YEL	0.5 / 0.10	43	80	15
PL4-GRE	1.2 / 0.15	1750	30	10
PL4-RED	0.8 / 0.23	27	45	12
PL5-YEL	1.5 / 0.14	300	25	10
PL6-RED	1.5 / 0.40	157	90	20
PL6-YEL	2.5 / 0.23	942	90	30
PL7-GRE	0.44 / 0.09	400	5	8
PL7-RED	1.8 / 0.43	144	80	40
PL7-YEL	1.5 / 0.07	37	50	10
PL8-GRE	1.7 / 0.16	484	30	15
PL9-GRE	2 / 0.15	1280	28	10
PL9-YEL	3 / 0.41	240	60	40

Table S3. Summary of the RDA analyses showing the relative importance of the different components of explained variation of community composition in the abundant (100 OTUs - relative abundance ≥0.1%) and in the dominant fraction (12 OTUs - relative abundance ≥1%) for the coverage of 1000 sequences per sample (38 samples). The effects of each component (E – environment, S – space, U – urbanization at subplot and plot levels) are not (marginal effect) or are corrected (conditional effect) for the other components. NS means that a certain model did not explain the variation in community composition significantly, or that the % explained variance was not significant.

Model	Components studied							
	Marginal effect				Conditional effect			
	E	S	U(subplot)	U(plot)	E	S	U(subplot)	U(plot)
Abundant (107 OTUs)	0.184***	NS	NS	0.018*	0.167 ***	NS	NS	0.013^{NS}
Dominant (12 OTUs)	0.302***	NS	NS	0.057*	0.277***	NS	NS	0.032*

E – Environment, S – Space, U – Urbanization (NS – not significant, ***$p < 0.001$, *$p < 0.05$)

Table S4. Results of RDA analyses. Significant environmental variables identified by forward selection, explaining the variation (adjR^2) in bacterioplankton community composition. Environmental variables effects were not (marginal effect) or were corrected (conditional effect) for the abundant (100 OTUs - relative abundance ≥0.1%) and in the dominant fraction (12 OTUs - relative abundance ≥1%) for the coverage of 1000 sequences per sample (38 samples).

Model	Significant environmental variables			
	E (marginal effect)	adjR^2	E|S+U (conditional effect)	adjR^2
Abundant (107 OTUs)	pH	0.041***	pH	0.036***
	alkalinity	0.036**	alkalinity	0.038***
	conductivity	0.034**	dissolved organic carbon	0.018*
	total phosphorus	0.034**	total phosphorus	0.019*
	dissolved organic carbon	0.019*	abundance of *Daphnia*	0.017*
	abundance of *Daphnia*	0.020*		
Dominant (12 OTUs)	pH	0.078***	pH	0.039***
	abundance of *Daphnia*	0.051**	suspended matter	0.019**
	suspended matter	0.050**	total nitrogen	0.016*
	total nitrogen	0.032**	maximum depth	0.015*
	molybdenum (Mo)	0.031*	abundance of *Daphnia*	0.017**
	maximum depth	0.035*	molybdenum (Mo)	0.013*
	nickel (Ni)	0.025*		

E – Environment, E|S + U – Environment corrected for space and urbanization (plot and subplot)
(***$p < 0.001$, **$p < 0.01$, *$p < 0.05$)

Table S5. Results of the test for differences in community composition in the abundant fraction (100 OTUs - relative abundance ≥0.1%) for the coverage of 1000 sequences per sample (38 samples) at the plot and subplot levels as inferred from PERMANOVA analysis.

	SS	MS	F	R^2	*P*
U_{plot}	0.56	0.28	1.54	0.07	0.02
$U_{subplot}$	0.42	0.21	1.15	0.05	0.24
U_{plot}: $U_{subplot}$	0.87	0.21	1.20	0.12	0.16
Residuals	5.27	0.18		0.73	

Table S6. Results of the test for differences in community composition in the abundant fraction (100 OTUs - relative abundance ≥0.1%) at the plot and subplot levels as inferred from PERMANOVA analysis.

	SS	MS	F	R^2	*P*
U_{plot}	0.59	0.29	1.62	0.06	0.01
$U_{subplot}$	0.41	0.20	1.11	0.04	0.27
U_{plot}: $U_{subplot}$	0.79	0.19	1.07	0.08	0.29
Residuals	7.52	0.18		0.80	

Table S7. Results of the test for the response of bacterioplankton diversity in relation to urbanization at plot (U_{plot}) and subplot level ($U_{subplot}$).

Fixed effects	U_{plot}		$U_{subplot}$		U_{plot}: $U_{subplot}$	
Factor	F	P	F	P	F	P
Richness	$F_{2,19}$=0.79	0.46	$F_{2,22}$=0.55	0.58	$F_{4,22}$=0.14	0.96
Shannon Diversity	$F_{2,19}$=0.72	0.49	$F_{2,22}$=0.06	0.93	$F_{4,22}$=0.98	0.43
Chao1	$F_{2,19}$=0.09	0.91	$F_{2,22}$=0.13	0.87	$F_{4,22}$=0.40	0.80
ACE	$F_{2,19}$=0.38	0.68	$F_{2,22}$=0.33	0.71	$F_{4,22}$=0.48	0.75

Table S8. Results of the test for the response for the ten most abundant classes found in relation to urbanization at plot (U_{plot}) and subplot level ($U_{subplot}$).

Fixed effects	U_{plot}		$U_{subplot}$		U_{plot}: $U_{subplot}$	
Factor	F	P	F	P	F	P
Betaproteobacteria	$F_{2,19}$=0.76	0.47	$F_{2,22}$=0.0003	1.00	$F_{4,22}$=0.45	0.76
Actinobacteria	$F_{2,19}$=0.67	0.52	$F_{2,22}$=0.30	0.73	$F_{4,22}$=1.29	0.30
Bacteroidetes	$F_{2,19}$=0.39	0.67	$F_{2,22}$=0.10	0.89	$F_{4,22}$=1.53	0.22
Alphaproteobacteria	$F_{2,19}$=3.27	0.059	$F_{2,22}$=0.95	0.40	$F_{4,22}$=0.43	0.78
Unclassified	$F_{2,19}$=0.63	0.53	$F_{2,22}$=3.27	0.056	$F_{4,22}$=4.07	0.01

Gammaproteobacteria	$F_{2,19}$=0.82	0.45	$F_{2,22}$=1.64	0.21	$F_{4,22}$=1.52	0.22
Firmicutes	$F_{2,19}$=1.54	0.23	$F_{2,22}$=1.20	0.31	$F_{4,22}$=2.72	0.55
Epsilonproteobacteria	$F_{2,19}$=0.65	0.53	$F_{2,22}$=0.43	0.65	$F_{4,22}$=1.09	0.38
Unclassified (Proteo)	$F_{2,19}$=0.70	0.50	$F_{2,22}$=2.54	0.10	$F_{4,22}$=0.60	0.66
Deltaproteobacteria	$F_{2,19}$=0.85	0.44	$F_{2,22}$=0.05	0.94	$F_{4,22}$=1.32	0.29

Table S9. Results of the test for the response of 12 dominant OTUs (> 1% of local abundance) in relation to urbanization at plot (U_{plot}) and subplot level ($U_{subplot}$).

Fixed effects	U_{plot}		$U_{subplot}$		U_{plot}: $U_{subplot}$	
Factor	F	P	F	P	F	P
Agrococcus	$F_{2,19}$=1.09	0.35	$F_{2,22}$=0.79	0.46	$F_{4,22}$=0.53	0.71
unclassified_1	$F_{2,19}$=0.80	0.46	$F_{2,22}$=1.08	0.35	$F_{4,22}$=2.68	0.05
Limnohabitans_1	$F_{2,19}$=0.13	0.87	$F_{2,22}$=0.36	0.69	$F_{4,22}$=0.59	0.66
Polynuclobacter_1	$F_{2,19}$=0.56	0.57	$F_{2,22}$=1.49	0.24	$F_{4,22}$=0.34	0.84
Limnohabitans_2	$F_{2,19}$=0.39	0.68	$F_{2,22}$=1.63	0.21	$F_{4,22}$=0.64	0.63
Cytophagaceae	$F_{2,19}$=0.46	0.63	$F_{2,22}$=0.63	0.54	$F_{4,22}$=0.46	0.76
unclassified (Actinobacteria-class)	$F_{2,19}$=7.09	0.005	$F_{2,22}$=3.96	0.03	$F_{4,22}$=6.83	0.001
Polynuclobacter_2	$F_{2,19}$=0.48	0.62	$F_{2,22}$=2.27	0.12	$F_{4,22}$=1.24	0.32
Methylobacter	$F_{2,19}$=0.40	0.67	$F_{2,22}$=8.84	0.001	$F_{4,22}$=8.05	0.004
Acinetobacter	$F_{2,19}$=1.14	0.34	$F_{2,22}$=1.50	0.24	$F_{4,22}$=0.64	0.63
unclassified_2	$F_{2,19}$=0.84	0.44	$F_{2,22}$=2.38	0.11	$F_{4,22}$=3.62	0.02
Flavobacterium	$F_{2,19}$=0.68	0.51	$F_{2,22}$=0.45	0.63	$F_{4,22}$=0.45	0.77

Table S10. Results of the test for the response of the six environmental factors selected by forward selection driving bacterioplankton community composition in relation to urbanization at plot (U_{plot}) and subplot level ($U_{subplot}$).

Fixed effects	U_{plot}		$U_{subplot}$		U_{plot}: $U_{subplot}$	
Factor	F	P	F	P	F	P
pH	$F_{2,19}$=0.75	0.48	$F_{2,22}$=4.4	0.02	$F_{4,22}$=2.62	0.06
Alkalinity	$F_{2,19}$=3.15	0.06	$F_{2,22}$=0.78	0.46	$F_{4,22}$=0.56	0.69
Conductivity	$F_{2,19}$=2.33	0.12	$F_{2,22}$=1.11	0.34	$F_{4,22}$=0.65	0.63
Abundance of *Daphnia*	$F_{2,19}$=0.36	0.69	$F_{2,22}$=1.98	0.16	$F_{4,22}$=1.18	0.34
Total phosphorus	$F_{2,19}$=1.05	0.36	$F_{2,22}$=2.17	0.13	$F_{4,22}$=0.08	0.98

Copper concentration	$F_{2,19}$=0.47	0.62	$F_{2,22}$=2.10	0.14	$F_{4,22}$=0.11	0.97

Table S11. Values of the environmental variables explaining variation in bacterioplankton community composition selected by forward selection for the 50 sites sampled.

	Selected variables					
Pond codes	Conductivity (µS)	pH	Alkalinity (mg/L)	Total phosphorus (mg/L)	Abundance of *Daphnia* (ind/L)	Copper concentration
PL1-YEL	117.5	8.22	170	0.04559771	0	0.01
PL10-GRE	546	8.2	285	0.41561286	0	0.09
PL10-RED	1166	7.85	225	0.09781234	40.3125	0.06
PL10-YEL	956	7.96	270	0.3346269	55.04	0.08
PL11-GRE	520	7.85	245	0.2519048	19.17083333	0.09
PL11-RED	522	7.84	245	0.16830728	6.45	0.11
PL11-YEL	788	8.04	390	0.69505283	114.1888889	0.05
PL12-GRE	387	6.89	155	0.34007148	0	0.01
PL12-RED	369	10.14	55	0.04406831	0	0.33
PL13-GRE	724	7.83	330	1.02733871	75.035	0.03
PL13-RED	11787	7.1	200	0.82602529	3.225	0.02
PL13-YEL	388	8.38	185	0.03955927	0	0.04
PL14-RED	103.8	9.88	115	0.74173638	0	0.02
PL14-YEL	1152	7.64	550	1.88222001	3.225	0
PL15-GRE	568	8.64	230	0.08615264	0.04673913	0.16
PL15-RED	1055	8.19	225	0.27429968	49.62916667	0.17
PL15-YEL	1072	7.92	385	2.21088384	193.2133333	0.06
PL16-GRE	167.2	7.63	75	0.18261595	157.38	0.12
PL16-RED	270	7	175	0.04509036	0	0.04
PL18-YEL	862	7.81	370	2.10057272	272.3333333	0.1
PL2-GRE	164.3	6.7	60	0.17170041	5.52857143	0.07
PL2-RED	659	7.27	160	0.12772228	0	0.07
PL2-YEL	492	7.48	200	0.08245156	59.81777778	0.07
PL23-GRE	29.7	6.55	75	0.3548875	27.48928571	0.02
PL23-RED	308	8.09	170	0.04785223	0.05	0.02
PL23-YEL	538	7.67	180	0.35106684	0	0.02
PL24-RED	290	8.91	140	0.05251156	0	0.33
PL24-YEL	214.6	7.65	175	0.75556042	0	0.04
PL25-RED	692	8	275	0.34752642	113.5916667	0.05

PL25-YEL	368	8.24	195	0.0355312	0.775	0.11
PL26-GRE	409	7.9	180	0.58058345	0	0.08
PL26-RED	127.2	7.16	265	0.07635301	0	0.09
PL26-YEL	164.8	7.55	115	1.93581623	345.075	0.16
PL27-GRE	309	7.53	180	0.87018378	0	0.28
PL27-RED	220	8.46	160	0.2369348	0	0.75
PL27-YEL	259	7.56	145	1.12912268	7.525	0.06
PL3-GRE	860	7.37	245	0.14274939	0.125	0.18
PL3-RED	463	7.47	200	1.86279623	1.775	0.03
PL3-YEL	479	8.02	145	1.62861323	405.6333333	0.05
PL4-GRE	275	6.83	70	0.23443612	0	0.06
PL4-RED	278	7.6	155	0.16613926	0.1	0.03
PL5-YEL	439	7.845	145	4.23545085	19.03382353	0.03
PL6-RED	751	7.7	305	0.03047741	31.7125	0.02
PL6-YEL	242	7.96	40	0.05386427	0	0.05
PL7-GRE	108.2	6.62	75	0.67373645	376.25	0.2
PL7-RED	112	9.39	45	0.55908369	0.125	0.61
PL7-YEL	133	6.35	60	1.24815754	104.8125	0.07
PL8-GRE	489	7.65	150	0.26064865	253.7	0.05
PL9-GRE	1277	7.77	385	0.20112938	0	0.05
PL9-YEL	677	7.73	385	0.07078819	0.35833333	0.05

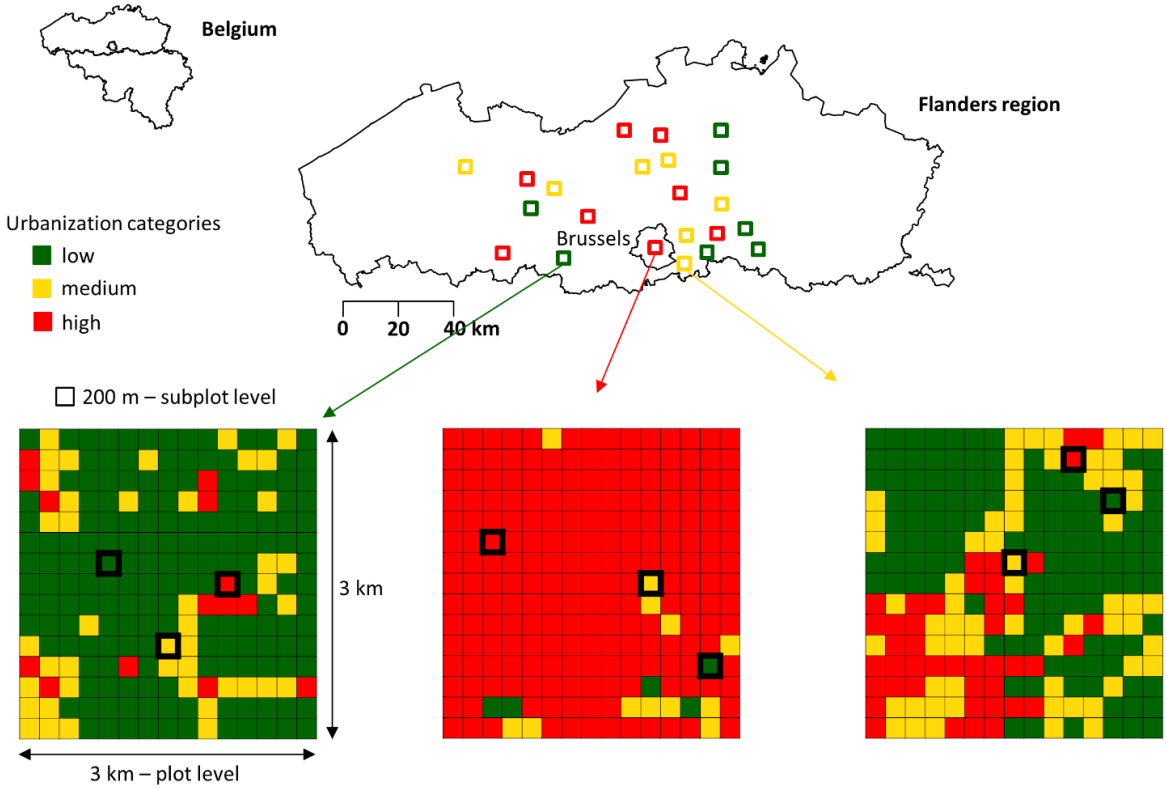

Figure S1. Localization of the Flanders region in Belgium and the representation of the hierarchical design; large squares represent 3 x 3 km plots, with grids of 200 m x 200 m subplots within plots, locations of ponds indicated as bold squares. Colours reflect urbanization categories: low (green), medium (yellow) and high (red). A detailed description of the hierarchical design is described in the main text. Maps of Belgium and Flanders region were produced in R using the package *BelgiumMaps.StatBel*.

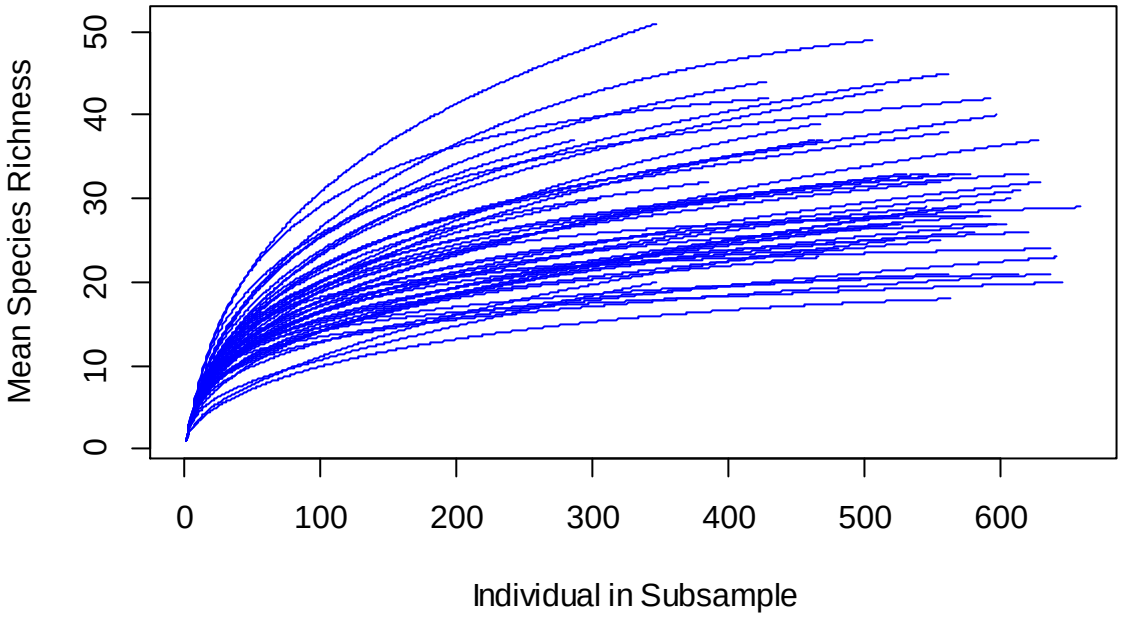

Figure S2. Rarefaction curves for each sample. Rarefaction curves are based on the normalized datasets (688 sequences/sample) with OTUs >0.1% of the total relative abundance (100 OTUs).

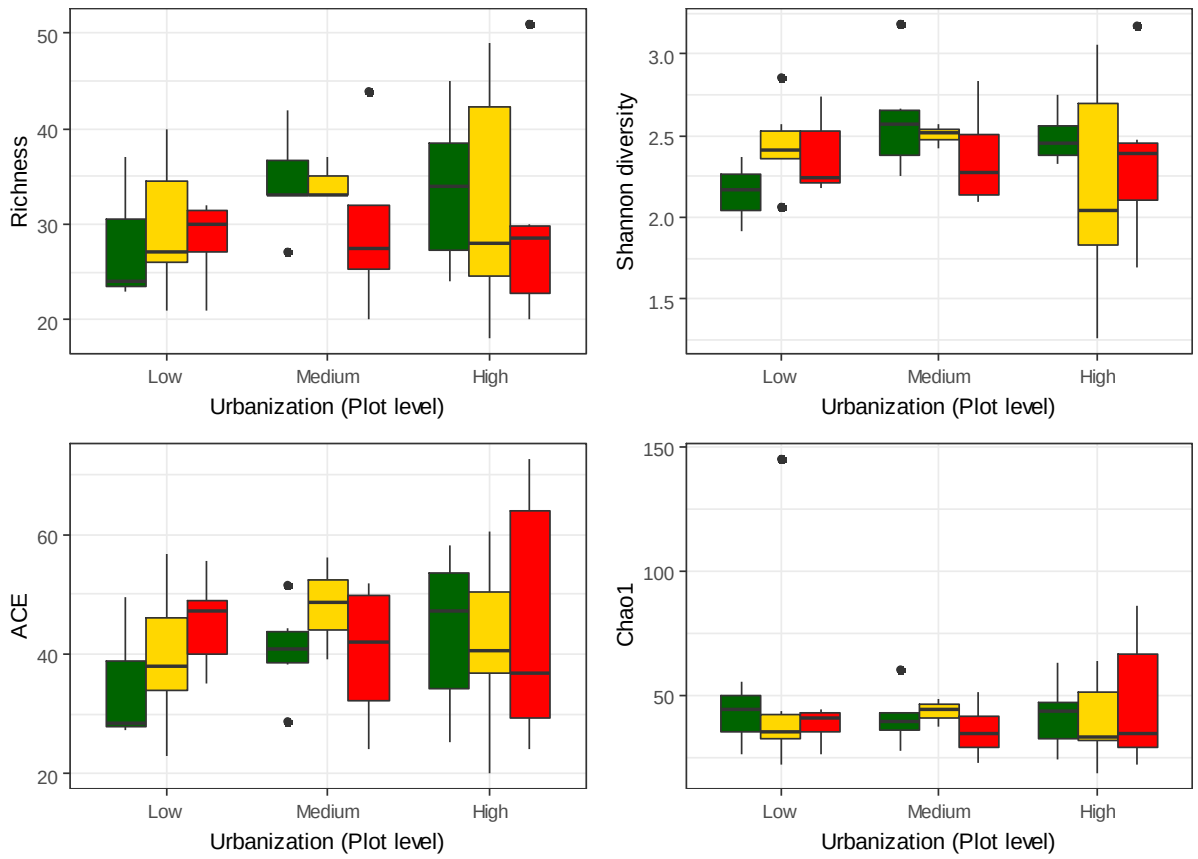

Figure S3. Boxplots of different diversity indices (Richness, Shannon, ACE and Chao1) along urbanization at subplot and plot level (x-axis). Urbanization categories at subplot level are represented by colours within the three categories at plot level (low, medium and high) depicted in the x-axis.

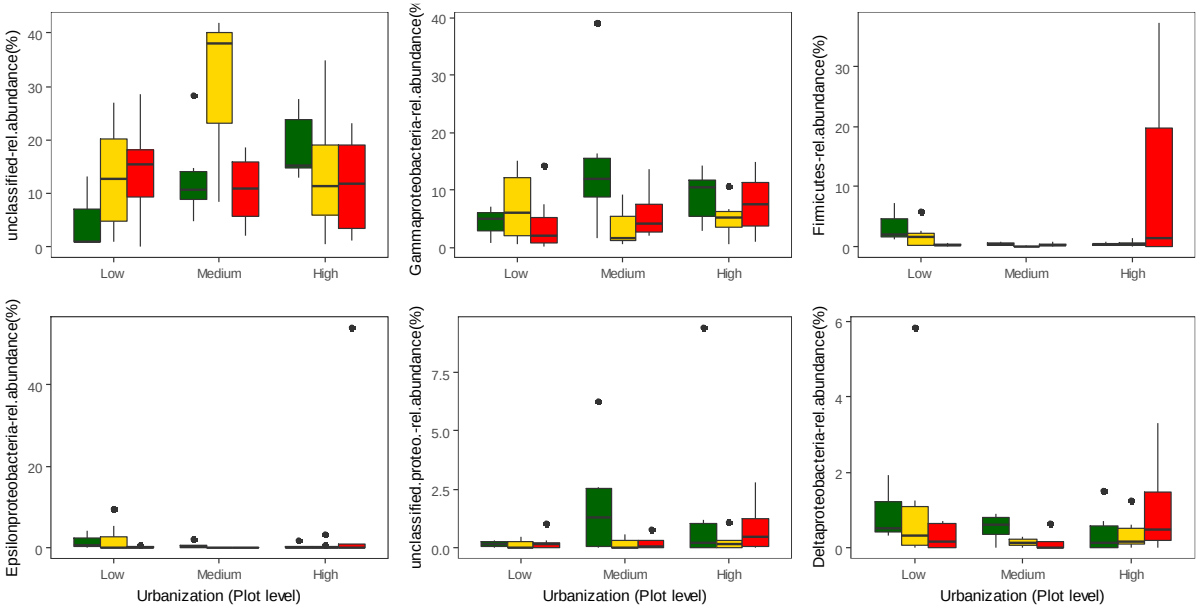

Figure S4. Boxplots of six abundant classes found along the urbanization gradient at subplot and plot levels. Urbanization categories at subplot level are represented by colours within the three categories at plot level (low, medium and high) depicted in the x-axis.

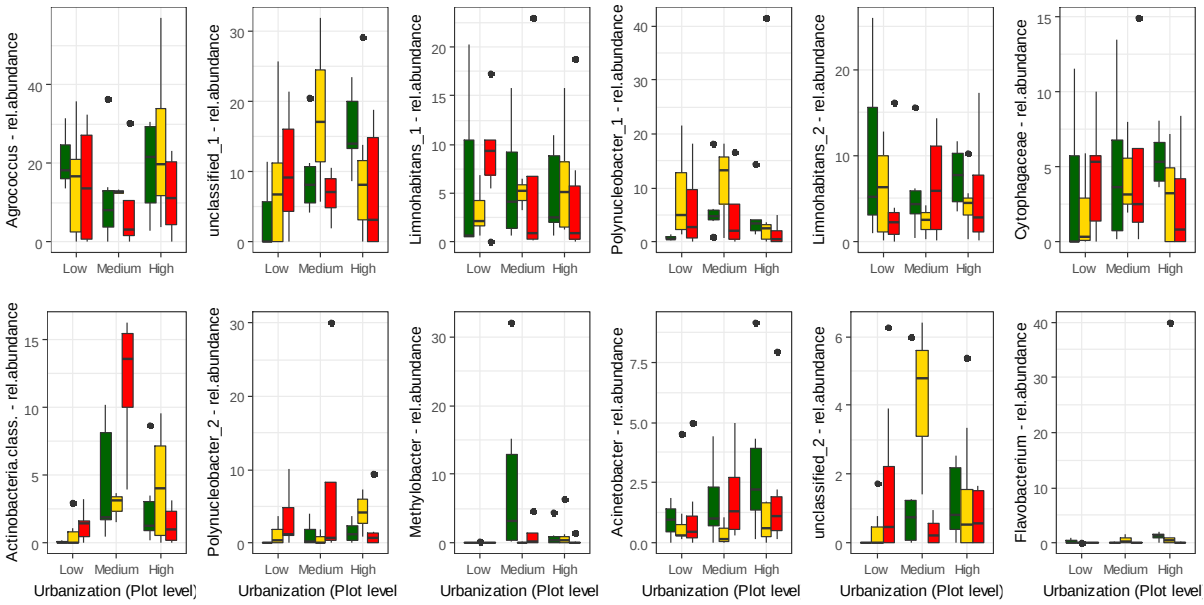

Figure S5. Boxplots of the 12 dominant OTUs (> 1% of local abundance) found along the urbanization gradient at subplot and plot levels. Urbanization categories at subplot level are represented by colours within the three categories at plot level (low, medium and high) depicted in the x-axis.

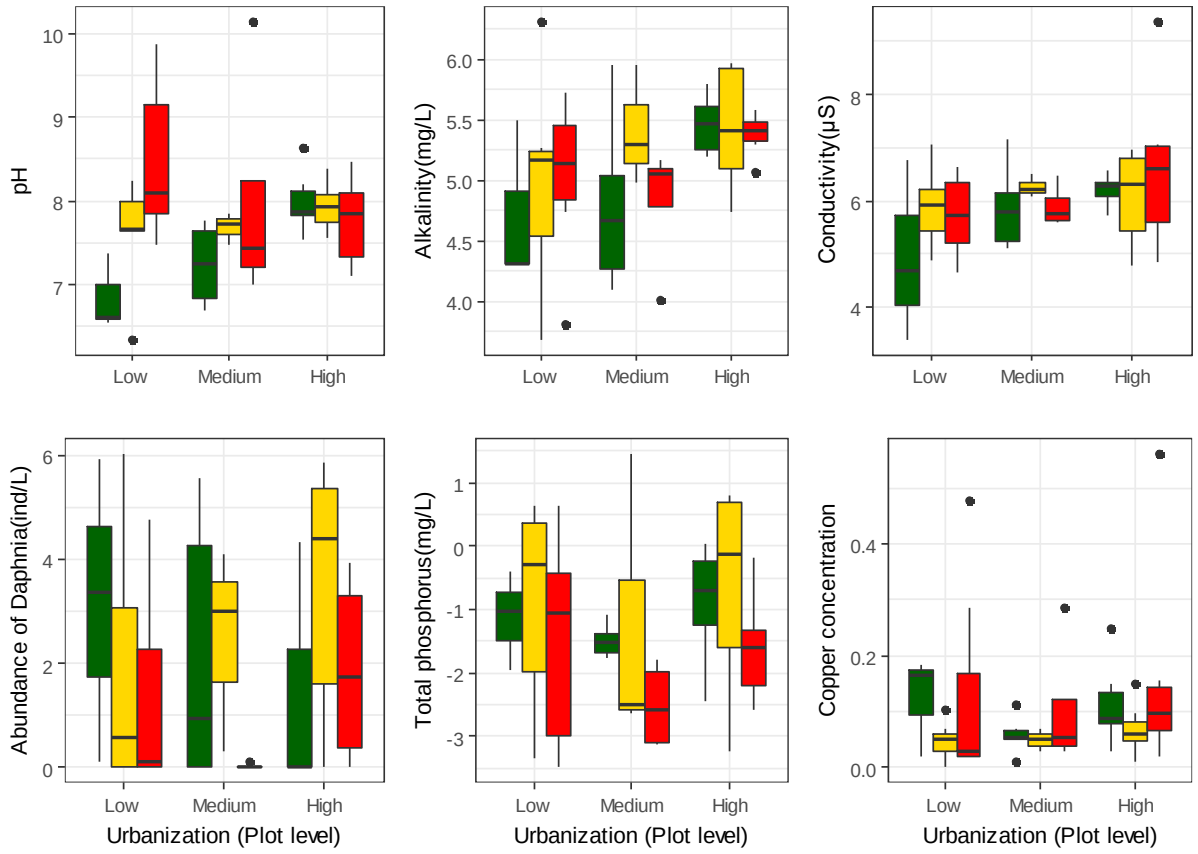

Figure S6. Boxplots showing the variation of the six environmental factors selected by forward selection driving bacterioplankton community composition among the urbanization categories at subplot and plot level. Urbanization categories at subplot level are represented by colours within the three categories at plot level (low, medium and high) depicted in the x-axis.

Chapter III

Bacterioplankton assembly along a eutrophication gradient is mainly structured by environmental filtering, including indirect effects of phytoplankton composition

Fabio Toshiro T. Hanashiro[1*], Luc De Meester[1,2,3], Matthias Vanhamel[1], Shinjini Mukherjee[1,4], Andros T. Gianuca[1,5], Laura Verbeek[6], Edwin Van den Berg[1] & Caroline Souffreau[1]

[1]Laboratory of Aquatic Ecology, Evolution & Conservation, KU Leuven, Charles Deberiotstraat 32, 3000, Leuven, Belgium

[2]Leibniz Institut für Gewässerökologie und Binnenfischerei (IGB), Müggelseedamm 310, 12587 Berlin, Germany

[3]Institute of Biology, Freie Universität Berlin, Königin-Luise-Strasse 1-3, 14195 Berlin, Germany

[4]Laboratory of Reproductive Genomics, KU Leuven, ON I Herestraat 49, 3000, Leuven, Belgium

[5]Departamento de Ecologia, Universidade Federal do Rio Grande do Norte, 59078-900, Natal, Rio Grande do Norte, Brazil

[6]Institute for Chemistry and Biology of the Marine Environment, University of Oldenburg, Schleusenstrasse 1, 26382 Wilhelmshaven, Germany

Modified from *Microbial Ecology* 2022, doi:10.1007/s00248-022-01994-x

3.1 Abstract

Biotic interactions are suggested to be key factors structuring bacterioplankton community assembly but are rarely included in metacommunity studies. Eutrophication of ponds and lakes provide a useful opportunity to evaluate how bacterioplankton assembly is affected by specific environmental conditions, especially also by biotic interactions with other trophic levels such as phytoplankton and zooplankton. Here, we evaluated the importance of deterministic and stochastic processes on bacterioplankton community assembly in 35 shallow ponds along a eutrophication gradient in Belgium and assessed the direct and indirect effects of phytoplankton and zooplankton community variation on bacterioplankton assembly through a path analysis and network analysis. Environmental filtering by abiotic factors (suspended matter concentration and pH) explained the largest part of the bacterioplankton community variation. Phytoplankton community structure was related to bacterioplankton structure through variation in chlorophyll-a and suspended matter concentration. Bacterioplankton communities were also spatially structured through pH. Overall, our results indicate that environmental variation is a key component driving bacterioplankton assembly along a eutrophication gradient, and that indirect biotic interactions can also be important in explaining bacterioplankton community composition. Furthermore, eutrophication was related to divergence in community structure and more eutrophic ponds had higher diversity of bacteria.

Keywords: bacterioplankton, eutrophication, community assembly, biotic factors

3.2 Introduction

Over the last decades, research has shown that microbial community assembly is driven by both deterministic and stochastic processes in soil (Caruso et al., 2011), aquatic habitats (Walker & Cyr, 2007; Chen et al., 2017), and the gut microbiome (Jeraldo et al., 2012). While stochastic processes such as ecological drift and random dispersal (Zhou & Ning, 2017) have only recently been reported to be important, it is already longer known that deterministic processes such as species sorting (Langenheder & Lindström, 2019) are structuring microbial communities. Species sorting, the process in which a species' presence and relative abundance are determined by the local environmental conditions, is often revealed as a main driver of microbial community structure (Van Der Gucht et al., 2007; Soininen, 2014; Monard et al., 2016).

Most studies testing for species sorting, however, only incorporate abiotic environmental variables in their analyses, such as water chemistry (Zwirglmaier et al., 2015), soil conditions (Lauber et al., 2013) or land use (Wang et al., 2011). Biotic interactions are rarely included as explanatory variables in bacterial community assembly (Van Der Gucht et al., 2007; Verreydt et al., 2012; Souffreau et al., 2015; Langenheder & Lindström, 2019). However, these types of interactions are crucial since they can structure microbial communities through predation and parasitism (e.g., grazing, viral infections) and host-associated microbiota (Langenheder & Lindström, 2019). So far, examples of the importance of biotic interactions in assembling microbial communities are limited to the top-down effects of predators on bacterioplankton (Chase et al., 2009; Souffreau et al., 2015), bacterial mortality related to prophage induction (Brum et al., 2016) and the ecological feedbacks of symbiotic relations between closely related communities [e.g., plants and belowground microbes (Hedlund et al., 2003), host development and gut microbiome (Burns et al., 2016)]. In freshwater lakes and ponds, observational studies assessing the relative importance of deterministic and stochastic processes along a latitudinal gradient (Souffreau et al., 2015) and an urbanization gradient (Hanashiro et al., 2019) have shown that densities of the zooplankton species *Daphnia* can explain a significant part of bacterioplankton assembly when included as biotic explanatory variable. Using microcosms with aquatic communities, Berga et al. (Berga et al., 2015) manipulated different environmental conditions, including the presence and absence of *Daphnia magna* and bacterioplankton dispersal rate, and demonstrated that

grazer presence combined with dispersal strengthened the relative importance of deterministic processes in bacterioplankton assembly. This could be explained by the dispersal of grazer-tolerant bacteria to all microcosms, decreasing idiosyncratic effects induced by efficient grazing by *D. magna*, which was itself dispersal limited (Berga et al., 2015). Biotic interactions can also act indirectly, with assembly processes cascading from higher trophic levels to the bacterial communities. Verreydt et al. (2012) studied how dispersal rates affected responses of three communities (bacterioplankton, phytoplankton and zooplankton) to a productivity gradient using mesocosms. Estimated by variation partitioning analysis, bacterioplankton assembly was mainly determined by indirect effects of dispersal limitation of *Daphnia*. Assessing and generalizing on the relative importance of deterministic and stochastic processes driving microbial community assembly would thus require an approach that also includes biotic interactions as explanatory variables. Moreover, it is urgent to better understand how these aquatic communities and their interactions are being impacted by the growing effects of human-induced processes on the environment (e.g., eutrophication, salinization and pollution).

In inland waters, eutrophication has been described as a strong anthropogenic pressure structuring bacterial communities (Wu et al., 2007; Zwirglmaier et al., 2015). Under eutrophic conditions, an increase in bacterial biomass is expected due to a higher availability of organic matter originating from phytoplankton growth (Özen et al., 2013). However, different responses to eutrophication in terms of structure and composition of bacterial communities have been reported. Heterotrophic bacterial communities showed a higher diversity and variability under nutrient-enriched conditions (Eiler & Bertilsson, 2004; Kiersztyn et al., 2019), while cyanobacterial communities tend to become more homogeneous (Monchamp et al., 2018). Nutrient enrichment of freshwater systems leads to phytoplankton proliferation (algae blooms) with often detrimental effects to other natural communities (Scheffer et al., 2001). The turbid state caused by eutrophication tends to favor Cyanobacteria over eukaryotic phytoplankton due to their adaptation to low light levels (Rigosi et al., 2014). Although higher phytoplankton densities could result in increased zooplankton grazing, toxicity of some cyanobacterial species has a negative effect on zooplankton and fish communities (Ger et al., 2014; Huisman et al., 2018). Additionally, oxygen-depletion due to an increase in bacterial heterotrophic activity can create anoxic conditions that are detrimental to fish (Vonlanthen et al., 2012). Despite the pivotal role of microbes in freshwater ecosystems

and increasing eutrophication levels, few studies evaluated the impacts of eutrophication on microbial communities and their interactions with other taxonomical groups at a metacommunity level. By combining a metacommunity perspective with the inclusion of biotic interactions across different taxonomical groups as putative explanatory variables, we can get a more integrated understanding on how regional and local environmental mechanisms generate and maintain local diversity under a eutrophication scenario.

This study aims to quantify the relative importance of deterministic and stochastic processes on bacterioplankton assembly along a eutrophication gradient. More specifically, we focus on the following questions: (i) Is bacterioplankton assembly mainly driven by environmental filtering in a eutrophication context? (ii) If yes, do abiotic or biotic variables explain most of community variation? (iii) Does eutrophication leads to more diverse bacterioplankton communities? To answer these questions, we sampled a bacterioplankton metacommunity of 35 shallow ponds in Belgium along a eutrophication gradient and combined different local abiotic environmental variables with different biotic variables (chlorophyll-a concentration, fish presence/absence, macrophyte infestation), including phytoplankton and zooplankton community composition. To assess whether and how bacterioplankton community assembly along a eutrophication gradient is structured by the local abiotic and biotic conditions, we combined a variation partitioning approach with a path analysis and network analysis.

3.3 Material and Methods

Study area and sampling strategy

We selected 35 shallow ponds along a gradient of eutrophication across Belgium (range of ~ 200 km of distance) (Fig. 1a). Pond selection followed a gradient of land use in the surrounding areas of the ponds (approximately within 100 m radius), ranging from intensive agriculture (croplands), extensive agriculture (pastures) and no agriculture (natural areas) (Declerck et al., 2006). The sampled ponds were the same as in the biodiversity analysis of phyto- and zooplankton carried out by Verbeek et al. (Verbeek et al., 2018). Sampling for the two studies was done simultaneously, so that we here can capitalize on the environmental and phyto- and zooplankton data collected by Verbeek et al. (Verbeek et al., 2018).

Sampling was carried out during Summer of 2013 (June-September), with each pond being sampled once. In the central area of the pond, a depth integrated water sample (total of 15 L) was collected using a pre-rinsed tube sampler (length 1 m; diameter 75 mm). In case of very shallow ponds (< 20 cm), a beaker (1 L) was used instead of the tube sampler. The water collected was pre-filtered using a 2 mm sieve to remove debris and larger materials. We used these water samples to characterize abiotic variables and bacterioplankton and phytoplankton communities. Bacterioplankton samples were then kept in 4 ºC in 2 L dark jars for further filtration in the laboratory on the same day of the sampling. We filtered up to 100 mL of pond water using MCE filters with 0.22 µm (Milipore® - U.S.A.), which were then kept frozen at -80 ºC until DNA extraction.

Characterization of the local abiotic and biotic factors

Local environmental factors (conductivity, pH, water temperature, total phosphorus and nitrogen, suspended matter, transparency, maximum depth, surface area, chlorophyll-a, macrophyte infestation, occupation by submerged, emergent and floating macrophytes and presence/absence of fish) were characterized using standardized protocols previously described in Declerck et al. (Declerck et al., 2006) and in De Bie et al. (De Bie et al., 2012).

To further evaluate the role of biotic interactions on the bacterioplankton community assembly, we also characterized phytoplankton and zooplankton community composition of the 35 ponds. A complete description of the methods used for the characterization of the local abiotic and biotic factors is available in the Supplementary Material.

DNA extraction and PCR of 16S rRNA genes

To characterize the bacterioplankton communities, DNA was extracted using the UltraCleanTM Soil DNA isolation kit (MoBio Laboratories Inc., Carlsbad, CA, U.S.A.) and amplified by PCR using the 16S rRNA universal eubacterial primers E338F (5'-ACTCCTACGGGAGGCAGCAGT-3') and E797R (5'-GGGTATCTAATCCTG-3') (amplicon length of 459-bp covering the variable V3 region of the 16S rRNA). Barcode sequences were inserted to distinguish samples. PCR mixtures were prepared as follows: 2.5 µl 10 x PCR buffer (Eurogentec, Seraing, Belgium), 1 µl $MgCl_2$ (50 mM; Eurogentec), 2.5 µl dNTPs (2 mM; Thermo Scientific, MA), 1.0 µl of each primer (20 pmol µl-1 each; Eurogentec), 0.2 µl Silverstar Taq DNA polymerase (5.0 U µl-1; Eurogentec) and 1 µl standardized template DNA (10.74 ng µl-1)

in a total reaction volume of 25 µL. PCR settings were as follows: initial denaturation at 94 ºC for 2 min, 30 cycles of 94 ºC for 1 min, 55 ºC for 1 min and 72 ºC for 1 min, and a final extension for 5 min (72 ºC). PCR quality was evaluated using gel electrophorese and purification was performed using a Qiaquick Gel extraction kit (Qiagen, Venlo, Netherlands). PCR products were quantified using PicoGreen dsDNA Assay Kit (Life Technologies, U.S.A.). Amplicons were sequenced in the Genomics Core of KU Leuven (Belgium) using a Roche 454 GS FLX+ (Switzerland).

Bioinformatic analyses

16S rRNA gene pyrosequences were processed using mothur (version 1.30.2) starting with the flowgram data. Following the barcode sequences, the sequences contained in the flowgrams (SFF file) were sorted per sample (pond). The default pipeline (http://www.mothur.org /wiki/Schloss_SOP) was followed for the analysis (Schloss et al., 2011). A 97% similarity level in the rRNA gene sequence was used for defining operational taxonomic units (OTUs). A total of 91,642 good-quality sequences were recovered and were available for downstream analyses. After quality filtering, 2321 OTUs (no global singletons) were identified from the 35 ponds. We rarefied samples to 1000 sequences per sample to enable comparisons among ponds. While sequencing depth was relatively low, 1000 sequences per sample has been shown to yield reliable beta-diversity estimations (Lundin et al., 2012) and almost a totality of samples reached saturation for the identified OTUs (Fig. S1).

Diversity and statistical analyses

All statistical analyses were performed in R (version 3.6.1). A complete description of the data analyses can be found in Supplementary Material. Briefly, to capture changes in bacterioplankton diversity along the eutrophication gradient, we tested for associations with alpha diversity metrics (number of OTUs, Shannon-Wiener index and ACE estimator) and the amount of suspended matter using linear and quadratic regressions. Suspended matter was used as proxy for eutrophication based on the PCA analysis and the forward selection (see Results). Changes in community composition (relative abundance data) were visualized in NMDS plots. PERMANOVA was performed to detect effects of eutrophication on community relative abundance data. To test for a distance-decay relation, the association of community similarity (abundance-based Bray-Curtis) with geographical distance and with environmental

heterogeneity were fitted each to a negative exponential model (Gómez-Rodríguez & Baselga, 2018).

To determine which environmental and spatial factors drive the variation in bacterioplankton community composition over the 35 ponds and quantify their relative importance, we performed a distance-based redundancy analysis (db-RDA). Spatial factors were represented by Moran's eigenvector maps (MEMs) (Dray et al., 2006) calculated based on the geographical coordinates of the ponds. Variation partitioning analysis was performed to quantify the relative contribution of environment and space on bacterioplankton community variation.

In order to evaluate possible causal relationships among spatial variables and standardized local environmental variables in determining the patterns of community composition in phyto- and zooplankton, and how all these co-determine bacterioplankton community composition, a path analysis (a variation of the structural equation modelling - SEM) was used. To complement our understanding of the bacterioplankton assembly at an individual level (OTUs for bacterioplankton, genus for phytoplankton and species for zooplankton), we calculated a correlation matrix including all response and explanatory variables used in the SEM model. Moreover, to visualize these correlations, we constructed a network representation. Detailed description of the network analysis is available in the Supplementary Material.

This combination of multiple statistical approaches was chosen to assess the different variables (environmental and spatial) driving the three distinct communities. Besides, some analyses were needed to reduce noise and for a better comparison among datasets with different orders of magnitude (e.g., bacterioplankton x zooplankton). Network and path analysis complemented the exploratory analyses addressing more specific relationships between/within the three communities and the selected factors influencing them.

3.4 Results

Pond environmental distribution along the eutrophication gradient

Principal component analysis (PCA) on the 12 environmental variables over the 35 studied ponds indicates that the first axis of variation (PC1, 28% of explained variation) is positively related to the concentration of nutrients (total phosphorus and total nitrogen), chlorophyll-a

and suspended matter (Fig. S2), which are all related to pond eutrophication. Pond area infested by macrophytes, fish presence/absence and pond area with floating and emerging macrophytes were all negatively correlated to the first PC axis (Fig. S2). Along the second axis (PC2, 14%), pH, submerged macrophytes, surface area and maximum depth of the ponds were positively correlated (Fig. S2). Chlorophyll-a and suspended matter were positively correlated (Pearson's correlation) to total phosphorus (Fig. S3a, Fig. S3b) and, total nitrogen (Fig. S3c, Fig. S3d) (chlorophyll-a: $r = 0.55$, $p < 0.001$and $r = 0.64$, $p < 0.001$, respectively; suspended matter: $r = 0.63$, $p < 0.001$ and $r = 0.72$, $p < 0.001$, respectively; all log-transformed). Since suspended matter was strongly related to proxies of eutrophication (total nutrients and chlorophyll-a concentrations), we used suspended matter values as a proxy for eutrophication for further analyses (Fig. 1a for a spatial distribution of the sampled ponds with respective suspended matter levels).

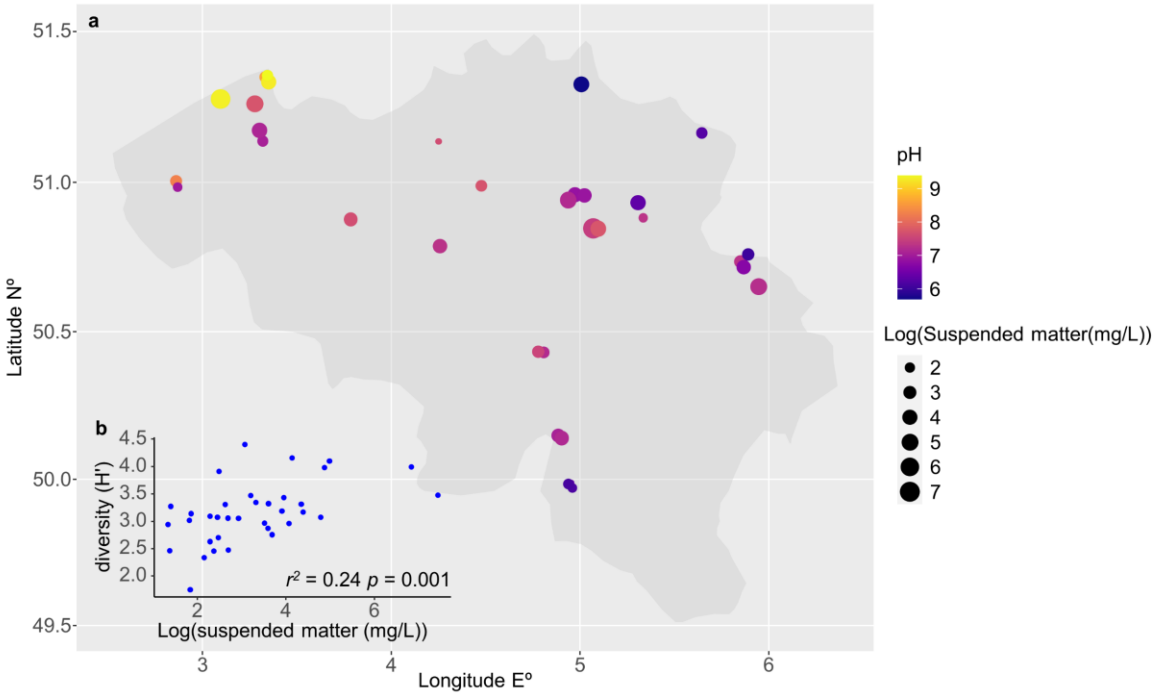

Figure 1. Map showing the distribution of ponds sampled across Belgium. (a) Ponds are coloured by pH values and circle sizes represent the log concentration of suspended matter. (b, inset) Linear regression of the Shannon-Wiener diversity index along the eutrophication gradient (suspended matter values used as a proxy for eutrophication).

Bacterioplankton diversity and composition along the eutrophication gradient

To assess the relationship between bacterioplankton diversity and the eutrophication gradient, we fitted linear and quadratic regressions. Linear models showed that concentration of suspended matter was positively related to both number of OTUs (R^2 = 0.14, p = 0.013) and Shannon's diversity (H') (R^2 = 0.24, p = 0.001) (Fig. S4a, Fig. S4b and Fig 1b). Linear models fitted better for both cases compared to the quadratic models in terms of adjusted-R^2. However, for the ACE estimator of bacterioplankton diversity (R^2 = 0.01, p = 0.235) (Fig. S4c), no significant relation was found.

A total of 19 phyla were found in the 35 sampled ponds and were present over the whole eutrophication gradient (suspended matter values) (Fig. S5). Proteobacteria was the most common phylum (mean relative abundance across samples ± standard deviation = 49.30 ± 8.92%) followed by Actinobacteria (22.16 ± 9.94%) and Bacteroidetes (14.37 ± 6.95%). A total of 11.56 ± 8.22% and 15.73 ± 10.50% of the OTUs could not reliably be assigned to a given phylum or class, respectively. At class level, 36 different class taxa were found with Betaproteobacteria (31.42 ± 11.14%), Gammaproteobacteria (8.20 ± 8.81%), Alphaproteobacteria (6.93 ± 7.01%) and Flavobacteria (5.52 ± 5.56%) being the more common classes (Fig. S6).

We also assessed the relationship between the most common phyla and classes and the amount of suspended matter using Pearson's correlation. Actinobacteria relative abundance at phylum level was negatively correlated with the amount of suspended matter (Table S1). With regard to the relation of suspended matter with the different trophic levels, Permanova indicated that there was a significant relation of suspended matter with the community structure of bacterioplankton (R^2 = 0.097, p < 0.001), phytoplankton (R^2 = 0.072, p < 0.001) and zooplankton (R^2 = 0.050 p = 0.011). An NMDS plot based on Bray-Curtis dissimilarity matrix (Fig. 2) revealed that the bacterioplankton communities tend to become more dissimilar with increasing eutrophication (Fig. 2a; NMDS – STRESS = 0.18). For phytoplankton communities, the first axis of variation (NMDS1) broadly aligned with the eutrophication gradient (Fig. 2b; NMDS – STRESS = 0.27). Zooplankton communities did not show a clear pattern, but some more dissimilar samples had higher values of suspended matter (Fig. 2c; NMDS – STRESS = 0.16).

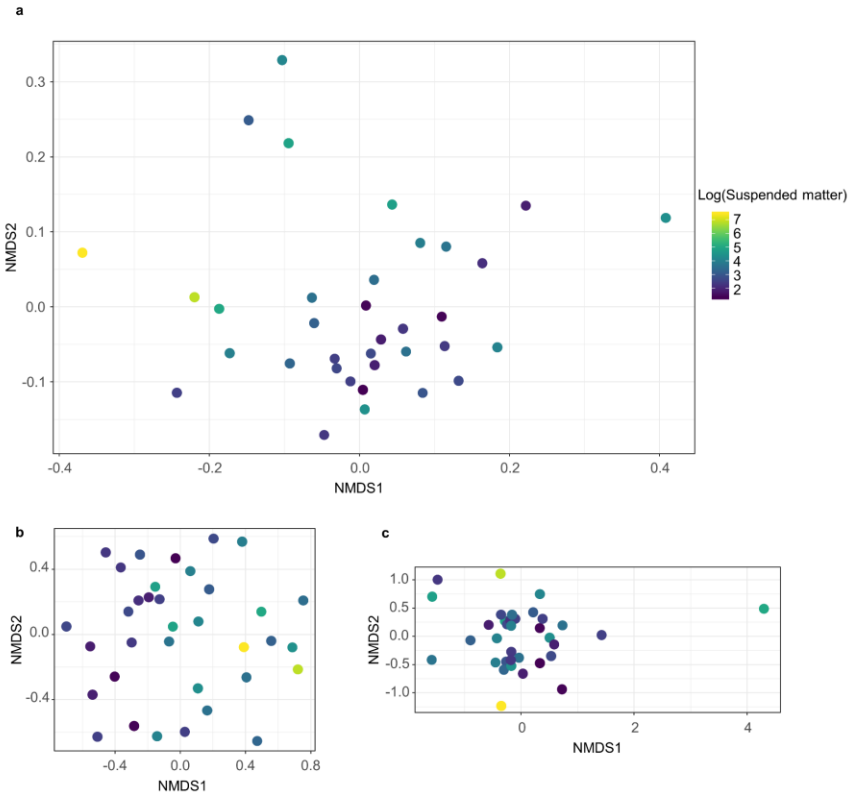

Figure 2. Non-metric multidimensional scaling (NMDS) plot of bacterioplankton (A), phytoplankton (B) and zooplankton (C) communities, colored by increasing concentration of suspended matter.

Relative importance of deterministic and stochastic factors driving bacterioplankton communities

Distance decay of bacterioplankton community similarity (Bray-Curtis distance) was relatively better explained by environmental heterogeneity (*pseudo* R^2 *= 0.082*, *p* = 0.001) (Fig. S7a) compared to geographical distance (*pseudo* R^2 *= 0.021*, *p* = 0.001) (Fig. S7b). Bacterioplankton community similarity decreased significantly in both cases, demonstrating that community variation could be at least partly attributed to the effect of space and environmental heterogeneity. As spatial distance was also positively correlated to environmental heterogeneity (Mantel test, r = 0.17, *p* = 0.004), the spatial relation in Fig. S7b might partly be an environmental effect.

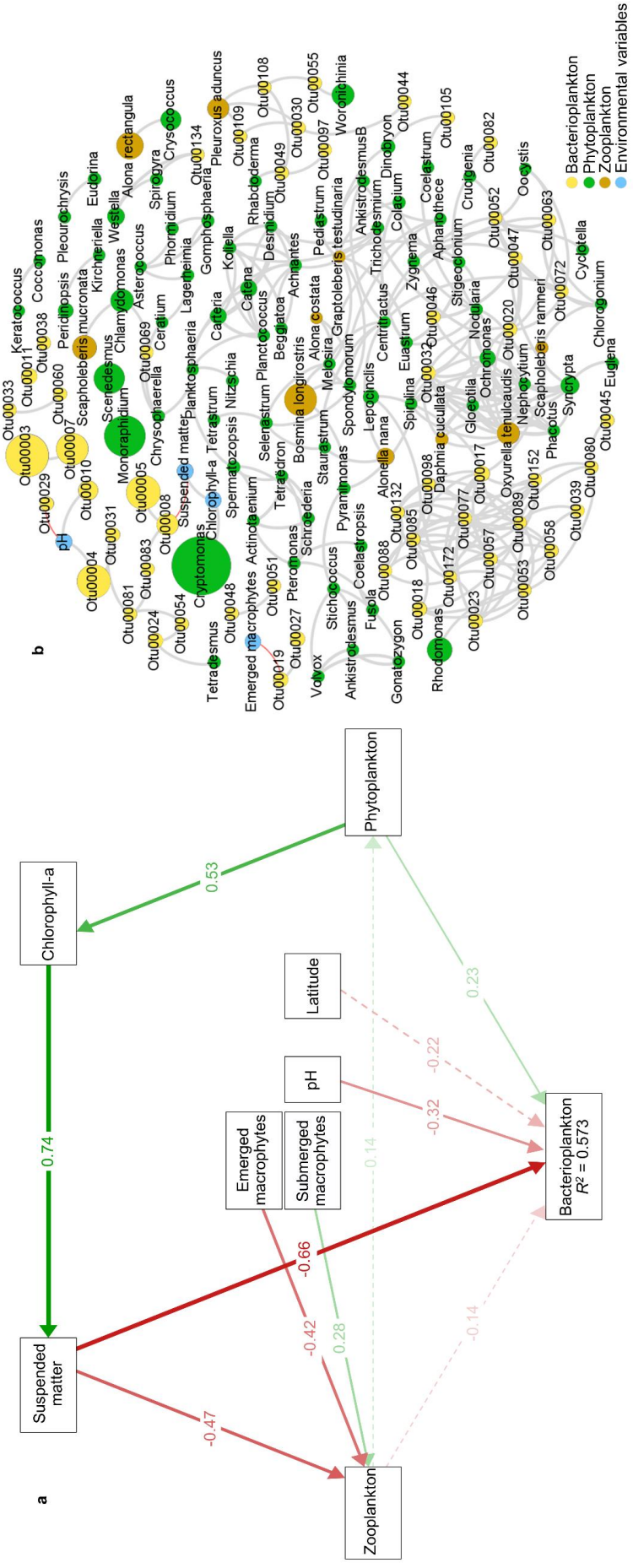

Figure 3. (A) Results of the best-fitting path analysis model. Thickness of lines are proportional to the path coefficients. Green and red arrows indicate positive and negative relationships, respectively. Dashed arrows represent not-significant relationships ($p > 0.05$). (B) Network of co-occurring bacterioplankton (yellow nodes), phytoplankton (green nodes), zooplankton (brown nodes) and environmental variables (blue nodes). Edges stand for Spearman's correlation values (> 0.6 – grey and < - 0.6 – red) ($p < 0.05$). Node's sizes are proportional to the taxa mean relative abundance across samples.

To assess which environmental and spatial variables are driving bacterioplankton, phytoplankton and zooplankton communities, and to quantify the pure and shared effects of both environment and space, we used a db-RDA approach. As response variables, we used the first PCoA eigenvector representing community variation for the three trophic levels bacterioplankton (global test – all environmental variables, F = 4.24, $adjR^2$ = 0.53), phytoplankton (F = 3.75, $adjR^2$ = 0.49) and zooplankton (F = 2.45, $adjR^2$ = 0.33) that maximized the amount of variation explained using all environmental variables. Since the global test was significant, we used forward selection procedure to assess which environmental variables were driving each specific community. Two environmental variables significantly explained bacterioplankton community variation (first PCoA axis), being suspended matter ($adjR^2$ = 0.37, p = 0.002) and pH ($adjR^2$ = 0.13, p = 0.008) (Table 1), while latitude ($adjR^2$ = 0.19, p = 0.004) was significant among the spatial variables driving changes in bacterioplankton assembly. For phytoplankton community variation, no variable was selected as an environmental variable significantly related to community variation. Zooplankton community variation was significantly explained by suspended matter ($adjR^2$ = 0.10, p = 0.046), % emergent macrophytes ($adjR^2$ = 0.12, p = 0.022) and % submerged macrophytes ($adjR^2$ = 0.06, p = 0.046) (Table 1). For both phytoplankton and zooplankton, no spatial variables were selected.

Table 1. Summary of the RDA analyses and the partial RDA analyses (variation partitioning) showing the explained variation of bacterioplankton, phytoplankton and zooplankton communities. NS = not significant (global test p > 0.05); [E] = environment; [S] = space; [E|S] = pure environment (controlled for space); [S|E] = pure space (controlled for environment); [E∩S] = shared effect by environment and space.

	[E]		[S]		[E|S]		[S|E]		[E∩S]	
	Var.	*p*	Var.	*p*	Var.	*p*	Var.	*p*	Var.	Variables selected
Bacterioplankton	50.8	0.001	19.9	0.004	33.3	0.001	2.3	0.118	18.0	Suspended matter pH Latitude
Phytoplankton	NS	-	NS	-	-	-	-	-	-	-
Zooplankton	29.7	0.002	NS	-	-	-	-	-	-	Suspended matter Emerged macrophytes Submerged macrophytes

Variation partitioning using the selected environmental and spatial variables showed that bacterioplankton community variation (first PCoA axis) was explained mainly by environment (50.8%) also when controlling for spatial effects [E|S] (33.3%). The linear spatial variable (latitude) explained 19.9% of community variation. Most part of this spatial signal was shared with the selected environmental variables [E∩S] (18%) (Table 1). Such a shared effect of environment and space most likely represents environmental variables that are spatially structured. To assess associations between significant environmental variables and latitude, we used Pearson's correlations. pH was positive associated with latitude (r = 0.43, p = 0.009) (Fig. 1a), while suspended matter did not have a significant association (r = 0.24, p = 0.149). Environment explained 22.4% of phytoplankton community variance (first PCoA axis), with space having no significant effect. Also, zooplankton community variance (first PCoA axis) was only explained by environmental variables (29.7%) (Table 1).

To investigate the role of biotic interactions on bacterioplankton community assembly, we performed a path analysis. We first used the results of the db-RDA described above to identify PCoA eigenvectors representing bacterioplankton, phytoplankton and zooplankton community variation. These eigenvectors and the environmental and geographical variables previously selected (Table 1) were combined in a model that included the possible relationships among the variables based on expert knowledge (Fig. S8). Our final model presented a good fit to the observed data (CFI = 1, p = 0.719) (Fig. 3a). The strongest factor directly affecting bacterioplankton community variation was suspended matter concentration (path coefficient = -0.66). Phytoplankton community variation (path coefficient = 0.23) influenced bacterioplankton variation positively, while zooplankton community variation did not have a strong direct effect driving bacterioplankton (path coefficient = -0.14) and phytoplankton variation (path coefficient = 0.14).

Phytoplankton community variation mediated suspended matter content through chlorophyll-a concentration (path coefficients 0.53 and 0.74). Zooplankton community variation was mediated by a combination of suspended matter (path coefficient -0.47), submerged macrophytes (0.28) and emerged macrophytes (-0.42) (Fig. 3a). In terms of the network analysis, a total of 142 nodes were correlated (in total 271 edges), with almost the totality of these correlations being positive (only 3 negative co-exclusions were detected) (Fig 3b). Environmental variables (pH, chlorophyll-a, suspended matter and emerged macrophytes) selected by the previous analyses were found significantly related to common

OTUs. Suspended matter was negatively correlated to different OTUs, but also positively related to chlorophyll-a and the most common phytoplankton taxa (*Cryptomonas*) (Fig. 3b).

3.5 Discussion

Here, we studied how environmental gradients and spatial drivers influenced bacterioplankton assembly of shallow ponds in Belgium along a eutrophication gradient. Our environmental variables did not only include gradients in abiotic conditions but also multiple biotic determinants. Overall, bacterioplankton assembly was mainly driven by environmental filtering (explained variation = 33.3%), with the concentration of suspended matter and pH being the most important variables. This observation is in line with previously reported results for bacterioplankton, where environmental filtering has regularly been observed to be the main determinant of bacterioplankton community variation (Logue & Lindström, 2010; Ren et al., 2015). This signal of environmental filtering is apparent in the significant distance-decay relationship that we observed, with communities inhabiting environmentally more dissimilar habitats being less similar. Our results also suggest that eutrophication levels can influence bacterioplankton assembly with communities responding strongly to specific eutrophic conditions. Higher levels of eutrophication in our study were associated to more dissimilar communities and may have selected for species that are more tolerant to the eutrophic conditions (e.g., having higher growth rates under eutrophic conditions thus becoming more competitive), reinforcing the relative importance of environmental filtering. Similar results were previously described in other microbial communities such as freshwater diatoms (Vilar et al., 2014) and coastal bacterioplankton communities (Dai et al., 2017), with environmental filtering being the most important assembly process along the eutrophication gradient. Here, the variation partitioning analysis showed that this environmental signal remained significant when corrected for spatial drivers and identified the total amount of suspended matter to be a key variable explaining variation in bacterioplankton community structure. In a multivariate analysis of all measured environmental variables, total suspended matter proved strongly associated with the concentration of total nitrogen, chlorophyll-a and total phosphorus along the axis of largest variation (PC1) (Fig. S2). Total suspended matter can thus be considered a good proxy of the overall eutrophication gradient that we aimed to sample. Additionally, ponds located in areas with intensive and extensive agriculture had the highest values of total

suspended matter, suggesting land use as a factor in the eutrophication context. pH, being related to the second axis of variation (PC2) in our multivariate analysis of environmental variables (Fig. S2), is known as an environmental variable that often is the driver of drastic changes in bacterial composition (Griffiths et al., 2011; Ren et al., 2015). The sampled ponds in our study did not present extreme pH values but ranged from 5.69 to 9.39 (mean = 7.26 ± 0.89), which could explain the secondary role of pH in structuring the sampled bacterioplankton communities. The effects of pH along a more continuous gradient with values ranging from 6.35 to 10.14 was previously described for microbial communities in shallow ponds in Belgium, in which pH explained the highest amount of community variation (Hanashiro et al., 2019).

Using a path analysis and co-occurrence analysis, we did not find a clear direct effect of phytoplankton or zooplankton community composition on bacterioplankton community variation, yet we found a strong indirect effect of phytoplankton community variation through changes in suspended matter via chlorophyll-a concentration (Fig. 3a and 3b). Phytoplankton communities are known to be able to elicit profound transformations in aquatic system due to their rapid growth under suitable conditions (Scheffer et al., 2001; Monchamp et al., 2018) and they can potentially directly influence bacterioplankton community composition through their excretion of organic compounds that can be used as resource by bacteria (Buchan et al., 2014). Our results thus suggest that eutrophication might have influenced bacterioplankton community composition by boosting algal growth. An increased phytoplankton density and/or higher productivity would result in a higher availability of dissolved organic carbon through the microbial loop and that way impact bacterioplankton community composition (Buchan et al., 2014). These nutrient-rich conditions might have favored certain bacterial groups (Kiersztyn et al., 2019). Another possibility is that changes in phytoplankton community composition would also impact the chlorophyll-a concentration. Specific phytoplankton taxa produce different levels of chlorophyll-a (Kasprzak et al., 2008) which would affect productivity levels and consequently bacterioplankton assembly. The negative effect of suspended matter on zooplankton community variation observed in the path analysis is in line with the observation that not all zooplankton species are able to cope with a high level of turbidity (Kirk, 1991) or an increase of phytoplankton growth in eutrophication scenarios (Monchamp et al., 2018). Zooplankton community variation, however, did not significantly influence bacterioplankton community variation, neither directly nor indirectly. The effect of

zooplankton on bacterioplankton communities is still in debate, with previous studies demonstrating the direct predator role of zooplankton on bacterial diversity (Berga et al., 2015), but also driving changes on community composition (Souffreau et al., 2015; Birtel & Matthews, 2016; Hanashiro et al., 2019).

In the present study, path analysis combined with the selection of variables by db-RDA proved to be an interesting approach to explore causal relations among a diverse set of abiotic and biotic variables, including community composition at different levels in the food web. Our choice to use PCoA eigenvectors to capture community variation as response variable instead of raw taxon-abundance data could have resulted in a reduced resolution. Direct signals of predation, symbiotic and host-parasite associations might be detected when analyzing species-specific connections (e.g. through co-occurrence networks) (Lima-Mendez et al., 2015) or mutually exclusive distributions (e.g. through checkerboard patterns) (Leibold & Mikkelson, 2002). Our network analysis confirms the results found by the variation partitioning and path analysis, with the same environmental variables correlating with the most common OTUs (pH – OTU00029, 00010 and 00004; suspended matter – OTU00008; emerged macrophytes – OTU00019) (Fig. 3b). A module (i.e., cluster of taxa that interact with themselves) with the most common OTUs was correlated with eutrophication levels, due to the link with suspended matter, chlorophyll-a and the most common phytoplankton taxon (*Cryptomonas*). Phytoplankton of the class Cryptophyceae have been described to have a high impact on the phytoplankton biomass in lakes because of their variable content in chlorophyll-a (Kasprzak et al., 2008). This indicates that phytoplankton identity can also contribute to the changes in environmental conditions related to eutrophication, subsequently affecting bacterioplankton assembly.

Spatial processes were also relevant for bacterioplankton assembly, but especially as shared variation with environmental gradients, likely caused by spatially structured variation in pH. While a simple isolation-by-distance correlation suggests a distance decay in which community similarity decreases with spatial distance (Fig. S7b), the variation partitioning analysis reveals that the pure spatial signal is not significant and that the overall spatial signal that is observed is due to a shared effect with environmental distance. This implies that the spatial signal is likely caused by a spatially structured environmental variable. The spatial driver that was selected by the variation partitioning analysis, latitude, was indeed found to be linked to variation in pH (Fig. 1a). This latitudinal gradient in pH reflects changes in soil

characteristics from highly productive and well buffered near the coast to more sandy, less buffered and more acidic in the north-east of the country (Gentile et al., 2009). Given that the pure spatial signal is not significant, our data suggest that the bacterioplankton community composition in the ponds studied is not significantly structured by dispersal limitation or mass effects, in line with earlier work on aquatic microbial communities (Van Der Gucht et al., 2007; Logue & Lindström, 2010). This interpretation does not preclude that dispersal limitation occurs (Bell, 2010), and one expects that for some of the less abundant or ecologically highly specialized taxa their occurrence might be influenced by dispersal limitation. Our result only indicates that, overall, dispersal limitation is not a key driver in structuring bacterioplankton communities in the habitats and landscape studied.

Approximately halve (47%) of bacterioplankton community variation in our study could not be explained by environment or space. This proportion of unexplained variation is relatively low compared to many other metacommunity studies (De Bie et al., 2012; Heino et al., 2014; Souffreau et al., 2015), and could be linked to many factors: unmeasured environmental variation, species interactions not captured by our analysis of phyto- and zooplankton community structure, unresolved spatial patterns, or stochasticity (Zhou & Ning, 2017; Langenheder & Lindström, 2019).

The communities we observed were dominated by the phyla Proteobacteria, Actinobacteria and Bacteroidetes, which are known to be typical freshwater bacteria of shallow ponds (Hanashiro et al., 2019; Kiersztyn et al., 2019). At the OTU level, more dissimilar communities were observed along the eutrophication gradient (Fig. 2a), yet no significant changes in the dominance of phyla or classes could be detected (Fig. S5 and Fig. S6). Similar observations were done in an analysis of the changes in bacterioplankton community structure in four Swedish lakes after a cyanobacterial bloom (Eiler & Bertilsson, 2004), in a comparative analysis of the changes in microbial communities of lakes characterized by different trophic state levels along different seasons (Zwirglmaier et al., 2015) and in interconnected lakes along gradients in nutrient enrichment (Kiersztyn et al., 2019). While eutrophication could lead to biotic homogenization (i.e. communities becoming more similar in community composition) due to the dominance of specific species under nutrient-rich conditions (Monchamp et al., 2018), eutrophication could also lead to divergence in processes structuring bacterioplankton communities. Such divergent structuring processes could include the occurrence of blooms of particular algae, changes from the clear-water to the turbid state

linked to the disappearance of macrophytes, strong increases in zooplankton grazing, or changes in the dominance of zooplankton grazers due to the proliferation of non-edible phytoplankton species (Scheffer et al., 2001; Ger et al., 2014).

In addition to the changes in community composition along the eutrophication gradient, we also observed an increase in diversity (number of OTUs and Shannon-Wiener index – Fig. S4a, Fig. 1b) with increasing eutrophication. Although the absolute values of alpha diversity might be underestimated due to the low sequencing-depth (rarefied to 1000 sequences per sample), the patterns of change are reliable (Lundin et al., 2012). The increase in alpha diversity together with more dissimilar communities along the eutrophication gradient (visualized by the NMDS – Fig. 2a and distance-decay relationships – Fig. S7b) suggest a biodiversity-productivity relationship in the studied ponds. Mechanisms underlying positive and hump-shaped relations between productivity and diversity have been heavily discussed (Smith, 2007). One hypothesis is that more productive environments would allow more species to reach sufficient densities to stably persist in a habitat. An increase in niche availability in productive environments might also weaken interspecific interactions and thus foster coexistence. High productivity can also, however, lead to increased dominance of one species (cfr. “blooms”) and in aquatic systems, the increase of productivity could lead to extreme conditions, such as large daily variation in the concentration of oxygen. Under such conditions, intermediate values of productivity might support more species (Waide et al., 1999). Despite the overall positive linear richness-eutrophication correlation, the highest number of OTUs (> 200) in single ponds were observed in ponds with intermediate productivity. We also observed higher zooplankton densities at intermediate values of eutrophication (Fig. S9), and zooplankton grazing can increase species diversity in bacterioplankton communities (Jürgens et al., 1994; Birtel & Matthews, 2016).

Here, we showed that bacterioplankton assembly in ponds along a eutrophication gradient was mainly governed by responses to environmental gradients that reflect differences in productivity. While we observed a spatial signal, this signal is non-significant when corrected for environmental drivers. Environmental filtering mainly involved responses to abiotic factors such as the amount of suspended solids and pH, but through a path analysis and a co-occurrence analysis we could also demonstrate an indirect role of phytoplankton community composition. Although we could not assess which factors were driving phytoplankton communities, our study shows the key role of phytoplankton under eutrophic

conditions, and its consequences for bacterioplankton assembly. We suggest that more studies on a wide array of systems should try to include data on community composition of other guilds in their analysis of microbial metacommunity structure, in an effort to get a better appreciation of the importance of food web and other species interactions in structuring microbial communities.

Acknowledgments

We thank Kristien Brans, Carla Denis, Jessie Engelen, Melissa Schepens and for their contributions to the field work and laboratory analyses. We also thank the different TippingPond partners at ISARA, University of Oldenburg and Uppsala University for their help with the landscape analyses and the contributions to the general development of the TippingPond project.

3.6 Supplementary material for Chapter III

Characterization of the local abiotic and biotic factors

We measured conductivity, pH and water temperature using a Hach® multimeter (United States) from water samples collected in the central area of each pond. Subsamples (4 x 50 mL) from the depth-integrated sample were taken for further analyses in the laboratory. Total phosphorus concentration was measured by the ascorbic acid method after perchlorate digestion (Murphy & Riley, 1962). Total nitrogen concentration was estimated using the Kjeldahl method (United States Environmental Protection Agency (USEPA), 1983). Suspended matter was measured gravimetrically in the laboratory by filtering a known volume of pond water (up to 1 L) on pre-weighted and dried GF/F filters (Whatman, GE Healthcare, U.K.). Filters were then dried again at 105°C for 6 h and the difference of weight between the pre-weighted and the filter with the suspended matter of the water sample was recorded as the suspended matter content. To estimate transparency, we used a Snell tube, an adaptation of the Secchi disk method. A dark gray plastic tube (length: 60 cm; diameter: 8 cm) is filled with the water sample and the depth at which a small 5 cm cross-section disk can be seen in the tube gives an estimation of transparency. This allows to measure transparency even in shallow systems where the Secchi depth would exceed water depth.

Morphometric characteristics of the ponds were also included as local factors. Maximum depth was measured in the center area of the pond using a graduated scale. Surface area of the ponds was estimated using the software Google Earth (version 7.1.8 – U.S.A.).

We included biotic parameters to our local environmental set. Chlorophyll-a concentration was measured with a fluorometer (Aquafluor® Turner Designs – U.S.A.). Five other biotic parameters were included: macrophyte infestation (visually estimated), occupation by submerged, emergent and floating macrophytes (visually estimated) and presence/absence of fish – determined with an electric fishing device following the Point Abundance Sampling approach (P.A.S.) (De Bie et al., 2012).

To characterize the phytoplankton community, a subsample (250 ml) of the depth-integrated water sample was taken and fixed with lugol's iodine solution for later characterization of phytoplankton community composition. Following the Utermöhl method, phytoplankton abundance was quantified as cell counts. Phytoplankton taxa were identified to genus level (Verbeek et al., 2018).

Zooplankton community composition was determined by collecting an additional depth-integrated water sample from four different locations in the pond using a tube sampler. These four water samples were pooled and subsequently filtered over a 4 mm sieve to remove plant material and other debris. 40 L was then poured through a conical zooplankton net (64 µm mesh size) and preserved with 4 ml of sugar saturated formalin. From each sample, at least 300 cladoceran and copepod individuals were counted using a stereomicroscope, but when a new species was found in the last 100 individuals, then the counting proceeded up to 400 individuals. Rotifers and ciliates were not included in our analyses. Cladoceran zooplankton identifications (species level for the genera *Alona*, *Bosmina, Chydorus, Daphnia, Graptoleberis, Oxyurella, Pleuroxus*, *Scapholeberis* and *Simocephalus;* genus level for *Ceriodaphnia*) were based on Flössner (Flössner, 2000). Copepods were classified into calanoids and cyclopoids.

Data analyses

First, a principal component analysis (PCA) was applied on the standardized environmental variables to visualize how ponds were distributed along the eutrophication gradient. All 14 environmental variables were tested for possible collinearity first (Pearson's correlation values > 0.8). Transparency and conductivity were removed from further analyses due to collinearity with suspended matter and pH, respectively. Subsequently, the 12 remaining environmental variables were log-transformed when necessary to achieve normality and standardized by z-score. To capture changes in bacterioplankton diversity along the eutrophication gradient, we tested for associations between alpha diversity (defined as number of OTUs, Shannon-Wiener index and ACE estimator) and the amount of suspended matter using linear and quadratic regressions. Suspended matter was used as proxy for eutrophication based on the PCA analysis (see Results).

Changes in community composition (relative abundance data) were visualized in NMDS (nonmetric multidimensional scaling) plots using Bray-Curtis dissimilarities. Prior to NMDS calculation, community data were square root transformed and submitted to a double Wisconsin standardization. Permutational Multivariate Analysis (Permanova) was performed to detect effects of eutrophication (suspended matter values) on community relative abundance data (number of permutations = 9999). To test for a distance-decay relation, the association of community similarity (abundance-based Bray-Curtis) with geographical

distance and with environmental heterogeneity (based on the Euclidean distance among the 12 environmental variables) were fitted each to a negative exponential model (Gómez-Rodríguez & Baselga, 2018). Since we used all pairwise combinations of sites (ponds) for these analyses, a randomizations test (1000 times) was used to estimate significance for the distance-decay models (Nekola & White, 1999).

To determine which environmental and spatial factors drive the variation in bacterioplankton community composition over the 35 ponds and quantify their relative importance, we performed a distance-based redundancy analysis (db-RDA) (Anderson, 1999). PCoA (Principal Coordinate Analysis) eigenvectors, orthogonal variables representing gradients in community composition variation, were used as response variables in the db-RDA (Gianuca et al., 2018) and were determined as follows. First, we calculated a distance matrix (Bray-Curtis dissimilarity) on the OTU relative abundance data. Then, pairwise values of the distance matrix were synthetized into PCoA eigenvectors (Anderson & Willis, 2003). Lingoes correction was applied when negative values were generated by the dissimilarity matrices (Borcard et al., 2011). All PCoA eigenvectors can in principle be used as response variables in the db-RDA analysis, as they can represent variation in community composition by unmeasured factors (Anderson & Willis, 2003). However, to avoid adding noise to the subsequent analyses, a selection procedure based on F-statistic and adjusted R^2 values was applied on the PCoA eigenvectors (Anderson & Willis, 2003; Duarte et al., 2012; Gianuca et al., 2018). This selection consists of performing the db-RDA analysis with sequential combinations of the PCoA eigenvectors (response variables) and all the environmental variables, adding an extra eigenvector every time (e.g., PCoA 1 ~ environmental variables; PCoA 1 + 2 ~ environmental variables). This procedure was done for all eigenvectors and the combination of eigenvectors that maximized F-statistic and adjusted R^2 values was selected as the final response variable used in the final db-RDA model.

After the selection of the PCoA eigenvectors as response variables for the three datasets (bacterioplankton, phytoplankton and zooplankton), an exploratory db-RDA global test using all 12 environmental variables as explanatory variables was performed for each of the taxonomical groups. Only if this global test was significant ($p < 0.05$), a forward selection procedure was applied (permutation test = 9999, $p < 0.05$) on the 12 standardized (z-score) variables (Blanchet et al., 2008).

Spatial factors were represented by Moran's eigenvector maps (MEMs) (Dray et al., 2006) calculated based on the geographical coordinates of the ponds. Original geographical coordinates were also included in the analysis to detect possible linear trends (Borcard et al., 2011).

Variation partitioning analysis was performed to quantify the relative contribution of environment and space on bacterioplankton community variation. The pure environmental component (E|S – environmental variation controlled by MEMs and geographical coordinates) and the pure spatial component (S|E – spatial variation is independent of any environmental factors) were assessed by a permutation test (999 permutations) (Peres-Neto et al., 2006). All community analyses described above were also performed on the phytoplankton and zooplankton relative abundance data (log-transformed). The choice of db-RDA rather the traditional redundancy analysis based on Hellinger-transformed data is justified by the possibility to compare results from datasets with different numbers of OTUs or species (Gianuca et al., 2018). This is important in the current study, as taxon richness differs across two orders of magnitude among the bacterioplankton, phytoplankton and zooplankton data sets. In addition, the PCoA eigenvectors can also be used as response variables in a path analysis (see below).

In order to evaluate possible causal relationships among spatial variables and standardized local environmental variables in determining the patterns of community composition in phyto- and zooplankton, and how all these co-determine bacterioplankton community composition, a path analysis (a variation of the structural equation modelling - SEM) was used. First, an a priori model was constructed based on expert knowledge on the causal relationships of variables affecting bacterioplankton community variation. The bacterioplankton first PCoA eigenvector was used as response variable in the model. Environmental and spatial factors that were previously selected by the db-RDA analysis and had a significant effect on bacterioplankton community variation were included in the model. Zooplankton and phytoplankton community variation (represented by the first PCoA eigenvector for both communities) were also incorporated into the model as biotic factors. In order to better understand the interaction among the three trophic levels, environmental variables related to zooplankton and phytoplankton communities were also included into the model. Model significance was tested by comparing the model-implied covariance matrix to

the observed covariance matrix with a Chi-square test (significant when $p > 0.05$) and by a Comparative Fit Index (> 0.90).

The correlation matrix was calculated using all possible pairwise Spearman's rank correlations (> 0.6 or < -0.6) between the most abundant OTUs (relative abundance ≥ 0.1%), phytoplankton, zooplankton and the environmental variables. We used a *p*-value of < 0.05 and to reduce the rate of false-positive results, *p*-values were corrected using the Benjamini-Hochberg method. To visualize the correlation matrix generated, we used a network representation built in the interactive platform Gephi (Bastian et al., 2009). Correlation matrix calculation and network visualization were developed following analyses described in (Ju et al., 2014).

R packages used in the data analyses: packages *vegan* (Oksanen, 2015), *ape* (Paradis et al., 2004), *betapart* (Baselga & Orme, 2012), *adespatial* (Dray et al., 2016), *gclus* (Hurley, 2019), *fossil* (Vavrek, 2011), *phyloseq* (McMurdie & Holmes, 2013), *ggplot2* (Wickham, 2009), *GGally* (Schloerke et al., 2018), *RColorBrewer* (Neuwirth, 2014), *lavaan* (Rosseel, 2012), *ggfortify* (Tang et al., 2016), *viridis* (Garnier, 2017) and *semPlot* (Epskamp & Stuber, 2017). Map of Belgium region was produced in R using the packages *maps* (Becker et al., 2018) and *BelgiumMaps.StatBel* (https://github.com/bnosac/BelgiumMaps.StatBel).

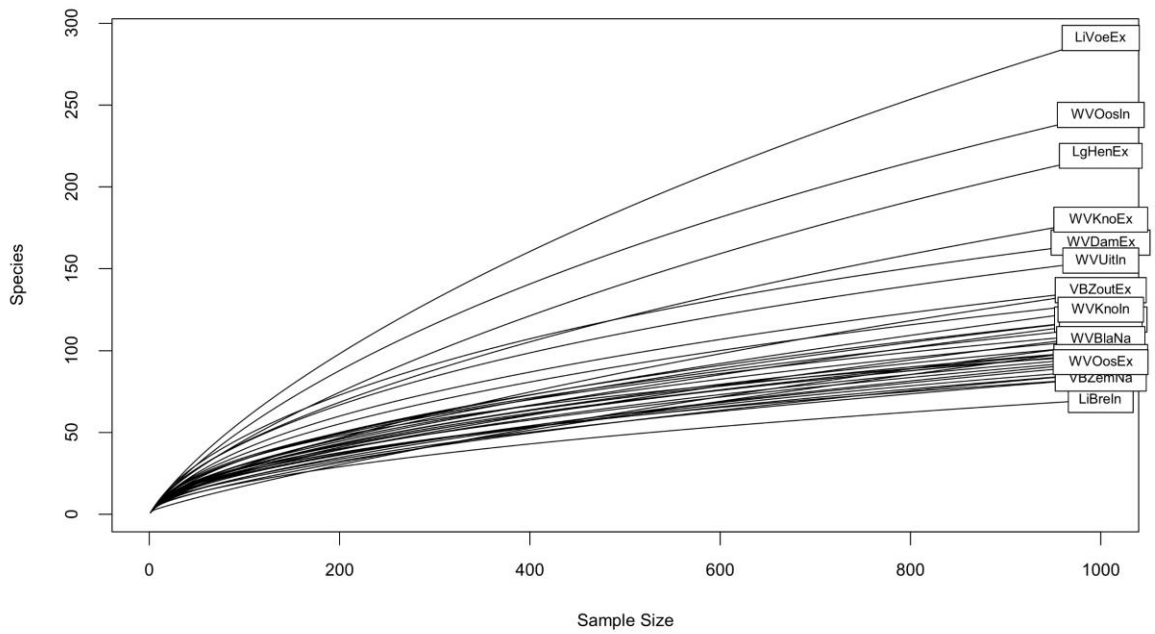

Figure S1. Rarefaction curves for each sample.

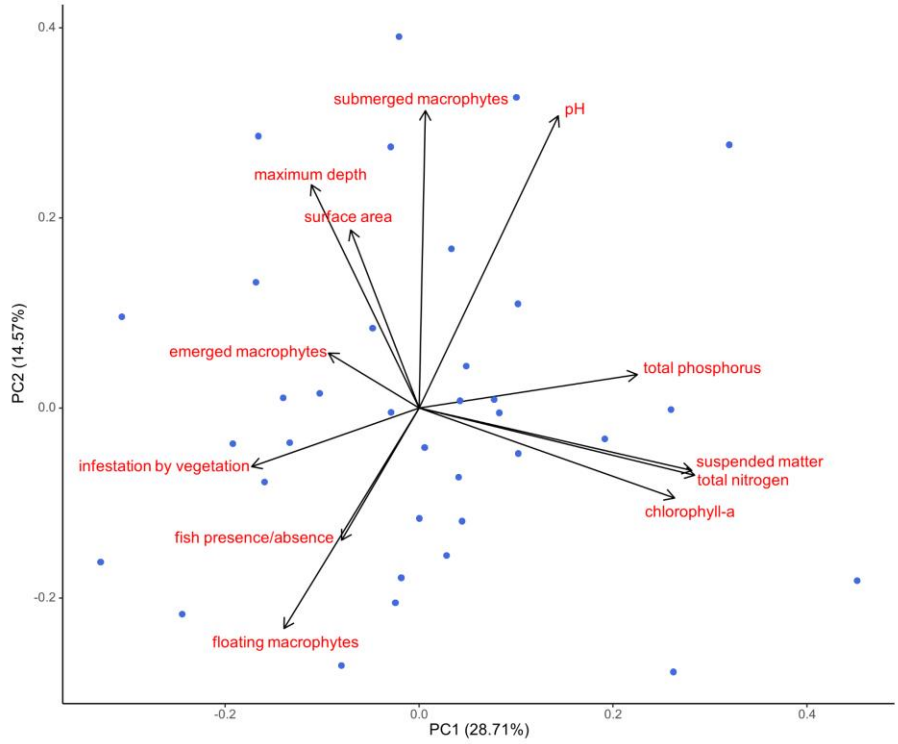

Figure S2. Principal component analysis (PCA) on the 12 standardized environmental variables. Blue dots represent the sampled ponds.

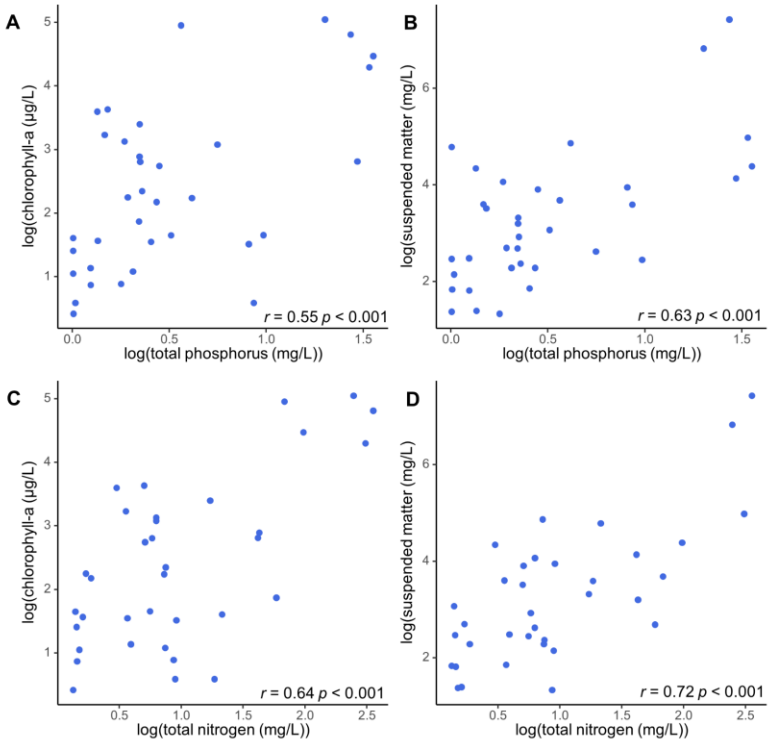

Figure S3. Correlation among chlorophyll-a and suspended matter with nutrients (total nitrogen and phosphorus) concentrations (all log-transformed). Pearson's correlation and significance values are depicted in each plot.

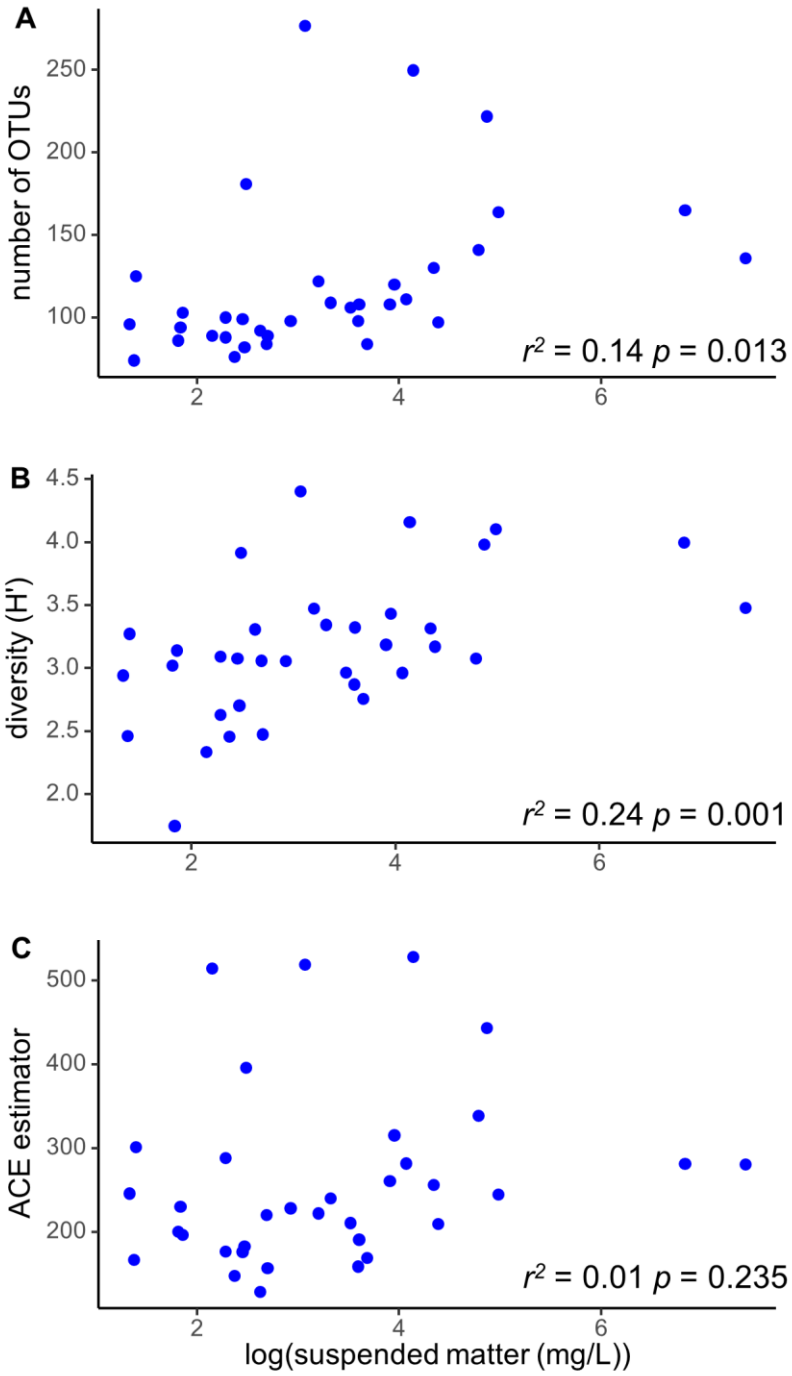

Figure S4. Linear regressions of different diversity indices (Number of OTUs, Shannon-Wiener diversity index and the ACE estimator) along the eutrophication gradient (suspended matter values used as proxy for eutrophication).

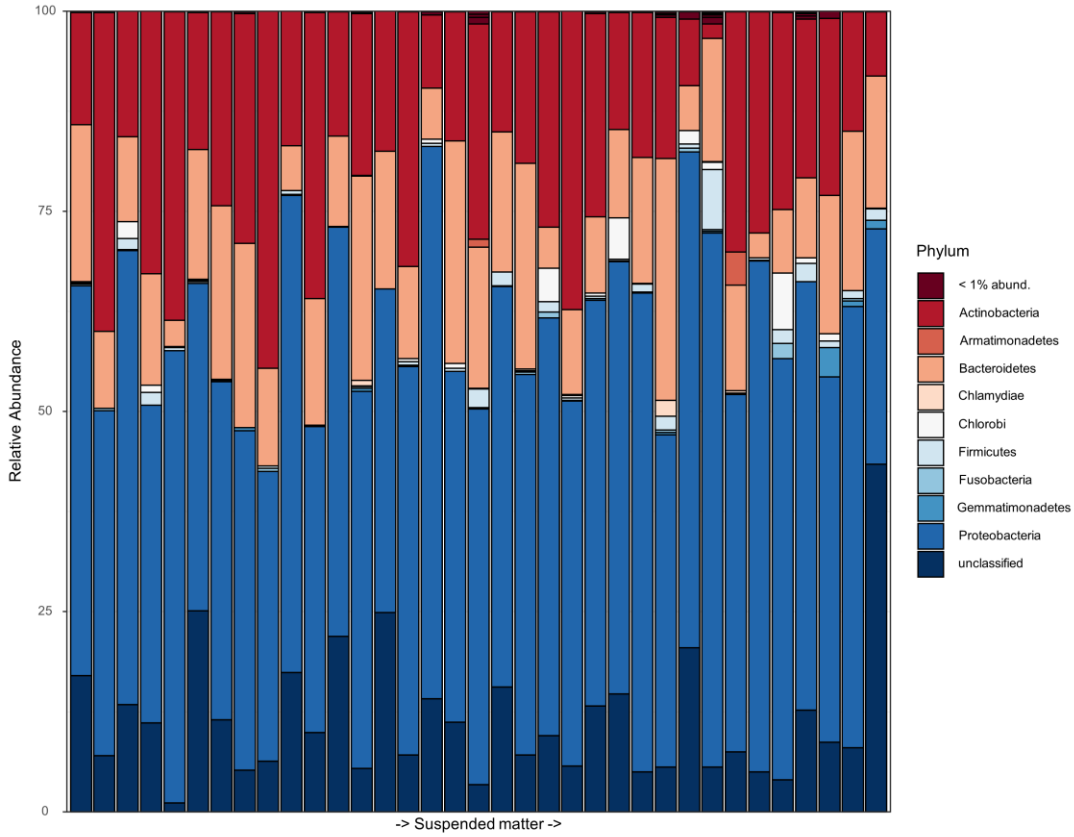

Figure S5. Relative abundance profiles at Phylum level in 35 shallow ponds in Belgium. Bars are ordered by increasing eutrophication (ordered by suspended matter values). Most common phyla were represented by Proteobacteria (49.30 ± 8.92%), Actinobacteria (22.16 ± 9.94%) and Bacteroidetes (14.37 ± 6.95%). Unclassified bacteria at phylum level accounted for 11.56 ± 8.22%.

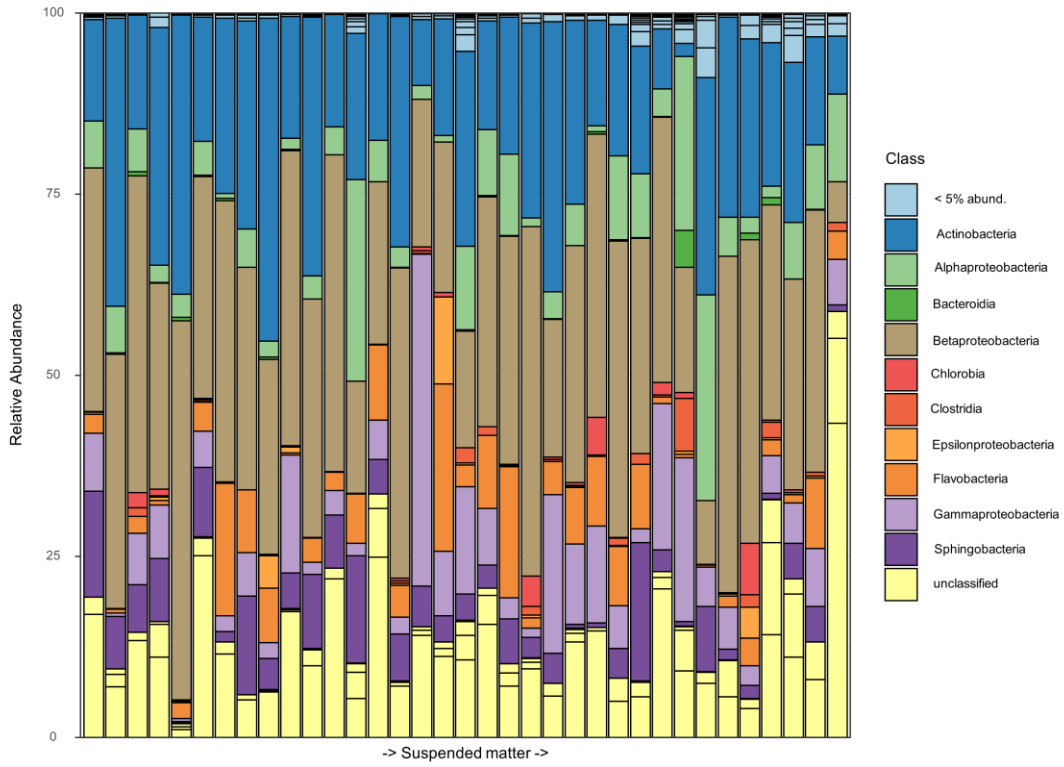

Figure S6. Relative abundance profiles at Class level in 35 shallow ponds in Belgium. Bars are ordered by increasing eutrophication (ordered by suspended matter values). Most common class were represented by Betaproteobacteria (31.42 ± 11.14%), Actinobacteria (22.16 ± 9.94%), Gammaproteobacteria (8.20 ± 8.81%), Alphaproteobacteria (6.93 ± 7.01%) and Flavobacteria (5.52 5.56%). Unclassified bacteria at phylum level accounted for 15.73 ± 10.50%.

Table S1. Pearson's correlation between the most common phyla and classes and the amount of suspended matter.

Taxa	Correlation value	*p*-value
Phylum level		
Actinobacteria	-0.34	0.044
Proteobacteria	0.05	0.764
Bacteroidetes	0.09	0.572
Class level		
Betaproteobacteria	-0.27	0.114
Gammaproteobacteria	0.06	0.704
Alphaproteobacteria	0.23	0.165
Flavobacteria	-0.004	0.979

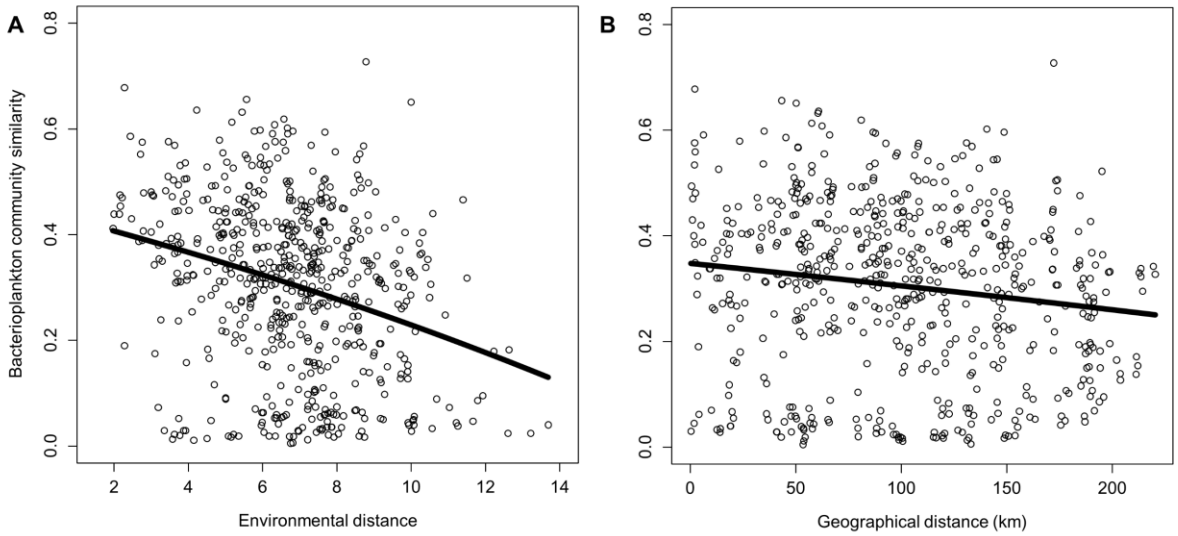

Figure S7. Relationships between bacterioplankton community similarity (Bray-Curtis distance), environmental distance (Euclidean distance) **a** and geographical distance **b** for all pair-wise comparisons of the 35 shallow ponds studied. Lines are drawn to indicate significant relationships ($p < 0.05$).

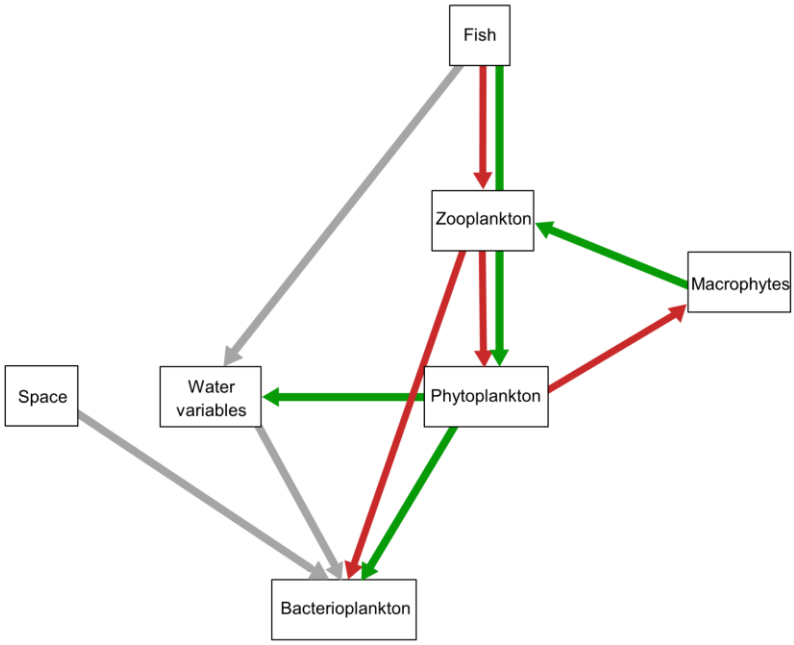

Figure S8. Proposed model for the path analysis showing the possible pathways among the environmental, spatial and biotic factors. Green and red arrows indicate positive and negative relationships, respectively. Grey arrows indicate relationships that could be either positive or negative. We omitted possible relationships among spatial and environmental variables with zooplankton and phytoplankton communities.

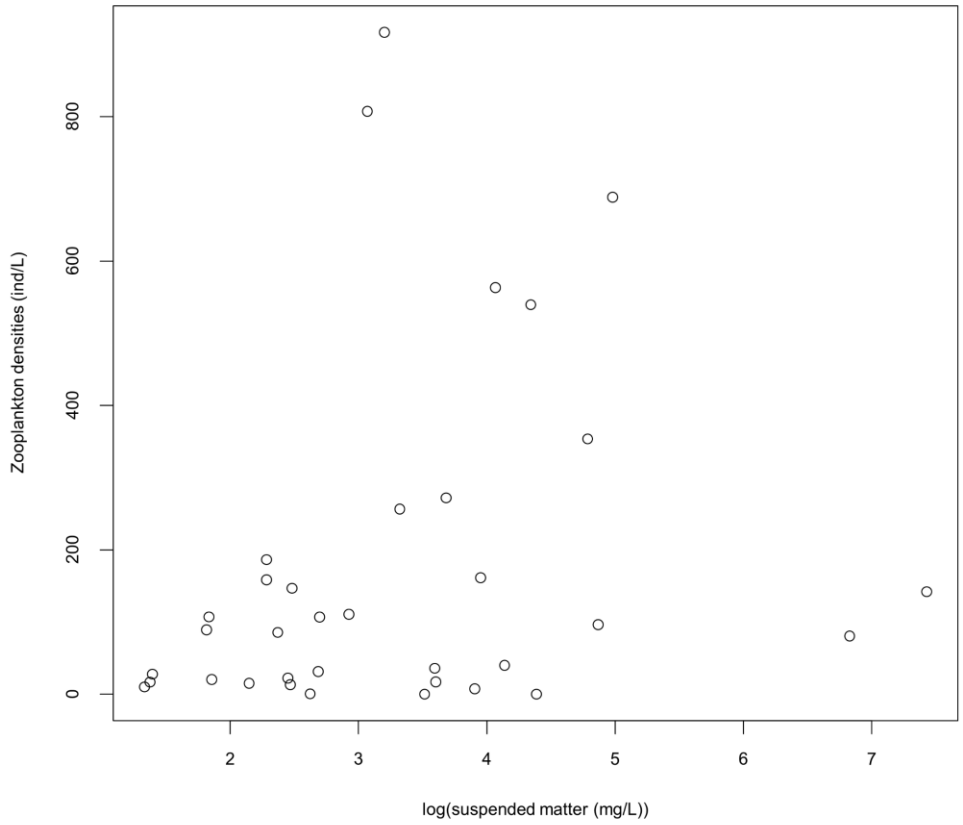

Figure S9. Zooplankton densities along the eutrophication gradient (suspended matter content values).

Chapter IV

Environmental filtering drives both taxonomic and trait structure in bacterioplankton communities of shallow ponds

Fabio Toshiro T. Hanashiro[1], Luc De Meester[1,2,3], Shinjini Mukherjee[4] & Caroline Souffreau[1]

[1]Laboratory of Aquatic Ecology, Evolution & Conservation, KU Leuven, Charles Deberiotstraat 32, B-3000 Leuven, Belgium

[2]Leibniz Institut für Gewässerökologie und Binnenfischerei (IGB), Müggelseedamm 310, 12587 Berlin, Germany

[3]Institute of Biology, Freie Universität Berlin, Königin-Luise-Strasse 1-3, 14195 Berlin, Germany

[4]Laboratory of Reproductive Genomics, KU Leuven, ON I Herestraat 49, B-3000 Leuven, Belgium

Publication status: in review in *Hydrobiologia*

4.1 Abstract

Within community ecology, the study of functional traits has been claimed to improve our understanding of the drivers of community variation. In microbial communities, the immense taxonomic diversity allows for much environmental filtering while also supporting functional redundancy. Here, we assess for 73 bacterioplankton communities inhabiting shallow ponds in Belgium the relative importance of environmental and spatial processes in assembling microbial communities at the level of species occurrence (taxonomy), at the level of functional group annotations, and at the level of carbon utilization profiles. Community structures as determined by functional group annotations and taxonomy were mainly driven by environment, yet with different biotic and abiotic variables being selected. Phosphorus concentration explained both functional and taxonomic community variation. Dissolved oxygen structured functional group annotations, while pH was relevant for taxonomic structure. Pure spatial variables explained little variation in functional group annotations and taxonomic structure. Variation in carbon utilization profiles could not be explained by environment nor space. Our results demonstrate that a combination of functional and taxonomic approaches provides complementary results to quantify the potential drivers of bacterioplankton assembly. The lack of explanatory power for carbon utilization profiles might reflect that our characterization of the environment did not include carbon source assessments.

Keywords: bacterioplankton, functional traits, shallow ponds, environmental filtering

4.1 Introduction

Improvements in the identification methods for microbial communities through, amongst others, next generation sequencing resulted in a microbial "revolution" and a much better view on the diversity of microorganisms (Prosser et al., 2007; Hug et al., 2016). The tremendous bacterial diversity estimated by these new methodologies also improved our knowledge on their biogeography (Hanson et al., 2012; Langenheder & Lindström, 2019; Martiny et al., 2006). Understanding the different processes by which bacterial communities are assembled across spatial scales is fundamental due to the pivotal role of bacteria in ecosystem functions such as biogeochemical cycles (Buchan et al., 2014). Overall, deterministic processes have been described as the major forces to determine bacterial distribution (O'Malley, 2007; Lindström & Langenheder, 2012). The main deterministic factor driving bacterial distribution is environmental filtering (i.e. the selection by local environmental conditions, also known as species sorting in metacommunity ecology) (Beisner et al., 2006; Langenheder et al., 2012; Logue & Lindström, 2010; Ren et al., 2015; Souffreau et al., 2015). However, it has been shown that other processes such as dispersal limitation, mass effects, historical contingency and ecological drift can also play an important role in bacterial assembly (Zhou et al., 2013; Rummens et al., 2018; Chen et al., 2019). Generally, a combination of both deterministic and stochastic processes along the community succession seems to be the best predictor of bacterial assembly (Langenheder & Lindström, 2019).

Besides the importance to understand the drivers of the taxonomic structure of bacterial communities, there is a growing interest in the patterns and drivers of their functional traits structure (Bissett et al., 2010; Krause et al., 2014; Lima et al., 2016; Louca et al., 2016; Ruiz-González et al., 2015). It has been argued that a trait-based approach could improve the interpretation of microbial community patterns because traits are directly related to the niches and functional roles of organisms in the ecosystem (Langenheder & Lindström, 2019; Martiny et al., 2013). Studies based on functional traits have suggested a high degree of functional redundancy among bacterial organisms, meaning that different bacteria can share similar functional traits (Böer et al., 2009; Comte & del Giorgio, 2010; Martiny et al., 2013). This functional redundancy could explain why trait-based and taxon-based community structure are explained by different mechanisms, especially in communities where thousands of species share similar functions. Environmental filtering has been shown as a relevant

process driving both functional (Louca et al., 2017) and taxonomic bacterial community structure (Crevecoeur et al., 2019), but they might be driven by different environmental variables. This has been shown for example for freshwater bodies in Canada in which the taxonomic bacterial composition was driven by pH, water temperature and water residence time, while the carbon utilization profiles of the communities were structured by the quality of dissolved organic matter (Ruiz-González et al., 2015). Globally across the oceans, bacterial and archaeal metabolic functional groups are closely related to local environmental conditions. Within each functional group, however, the link with environment is weak (Louca et al., 2016). In some soil and lake bacterial community studies, spatial factors (a proxy for different spatially-related processes such as dispersal limitation and mass effects, but often also including unmeasured spatially-structured environmental conditions) were more important than environmental variables in explaining the variation in functional genes and carbon utilization profiles (Shi et al., 2015; Lima et al., 2016). In yet another study on soil bacterial communities, Bissett et al. (2010) did not find a relation between carbon utilization profiles and environment or space, while for taxonomic composition both factors were important.

In the present study, we assessed the relative importance of environmental and spatial processes in structuring bacterioplankton taxonomic community composition (identified by 16S rRNA gene-amplicon sequencing) and bacterial functional traits characterized in terms of carbon utilization profiles and functional group annotations in 73 shallow ponds in Belgium. Carbon utilization profiles assess the metabolic potential for 31 specific carbon sources using a spectrophotometric laboratory assay, whereas functional group annotations categorize the genetically identified taxonomic taxa into broader functional groups (e.g., chemoheterotrophs) based on a taxonomic database. Because of functional redundancy, we therefore expect that functional groups will show a stronger match to environmental conditions (e.g., aerobic groups related to higher levels of dissolved oxygen) compared to taxonomic composition. We profited from an extensive field sampling along two different environmental gradients (eutrophication and urbanization), where we measured a broad range of abiotic and biotic environmental conditions in all sampled ponds. Thus, we could use a standardized set of local and regional variables in combination with variation partitioning analysis to disentangle patterns for taxonomic composition, carbon utilization and functional groups. Given that we assessed a broad array of abiotic and biotic environmental variables,

we expected to find a strong association between both taxonomic and functional diversity with environmental gradients yet might be driven by different environmental dimensions. Given that we did not assess dissolved and particulate organic carbon sources directly, we did not expect a to find a strong association between carbon utilization profiles and environmental gradients.

We also investigated if functional diversity was related to changes in taxonomic diversity. Due to high functional redundancy, we do not expect a linear relationship between functional and taxonomical diversity. Finally, based on the variables driving changes in taxonomic and functional bacterioplankton communities identified in this study, we assessed specific relations between individual taxonomic levels (phyla and classes) and individual metabolic functions with the environmental variables that were studied.

4.3 Methods

Selection of ponds and bacterioplankton sampling

We sampled 73 shallow ponds across Belgium from semi-natural to agricultural and urban areas (for map see Fig. S1). Ponds are excellent model systems in metacommunity assembly because they have diverse ecological characteristics, have a convenient patch structure in which local communities are very well delimited in space, and are influenced by changes in their immediate surroundings (De Meester et al., 2005). Bacterioplankton communities inhabiting these shallow ponds provide an extensive range of functional and taxonomical diversity. The field survey was carried out during summer of 2013 (May-September). We sampled bacterioplankton communities in the central area of each pond using a tube sampler (length 1 m, diameter 75 mm) or a beaker (ponds with less than 20 cm of depth). A total of 15 L of a depth-integrated water sample was collected and filtered through a 2 mm mesh size sieve to remove larger materials. This sieved water sample was used to characterize bacterioplankton taxonomic community composition and bacterioplankton functional traits (carbon utilization profiles). For the bacterioplankton taxonomic and functional characterization, 2 L of sieved pond water was kept at 4 ºC in the dark until further processing at the laboratory. For taxonomic characterization, up to 100 ml was filtered through a MCE filter with 0.22 µm pore size (Millipore) and the filter stored at -80 ºC until DNA extraction.

Measurement of abiotic and biotic explanatory variables

A total of 19 biotic and biotic factors were characterized following standardized protocols used in Declerck et al. (2006), Souffreau et al. (2018) and Hanashiro et al. (2019). A summarized description of the protocols is given here. Using a Hach multimeter (USA), we measured pH and conductivity from the depth-integrated water sample, while water temperature and dissolved oxygen were measured directly in the pond. To estimate water transparency, we used an adaptation of the classic Secchi disc method for shallow ponds, where a dark grey plastic tube (length 60 cm; diameter 8 cm) has a 5 cm Secchi disk disposed in the same fashion as the traditional version. Transparency was visually estimated as the depth that the disk could still be seen (De Bie et al., 2012). Total phosphorus and total nitrogen were quantified with the ascorbic acid method after perchlorate digestion (Murphy & Riley, 1962) and with the Kjeldahl method (United States Environmental Protection Agency (USEPA), 1983), respectively. Available phosphorus and nitrogen (dissolved nutrients) were determined after filtration of 50 mL of sampled water on a GF/F filter with a Technicon autoanalyzer III (United Kingdom). Suspended matter was quantified by filtering water samples (up to 1 L) using a pre-weighted GF/F filter (previously dried at 105 °C for 6 hours). Filters were then dried and weighted again and the difference of weight between the pre-weighted and the dried filter was used as suspended matter content.

Three morphometric variables were quantified. Surface area of the ponds was estimated in Google Earth (version 7.1.8, USA). The proportion of pond surface area that was shaded was visually estimated, and maximum depth was measured using a graduated scale.

In order to examine possible effects of other trophic levels on the bacterioplankton community, we included biotic factors related to phytoplankton and zooplankton communities. Chlorophyll-a concentration was estimated with a calibrated Aquafluor Turner Designs fluorometer (USA). Zooplankton communities were sampled using a tube sampler (length 1 m, diameter 75mm). A depth-integrated water sample (total of 60 L) at both pelagic and littoral zones from each pond was collected and filtered over a 2 mm mesh size sieve to remove larger debris. We then filtered a subsample (20 to 40 L) of the total volume over a 64 µm mesh size conical zooplankton net. Zooplankton identification was conducted using a stereomicroscope and a minimum of 300 individuals were identified; when a new species was found in the last 100 individuals, counting proceeded until 400 individuals. Zooplankton and bacterioplankton sampling were performed on the same day. More information about the

zooplankton communities and their identification can be found in Gianuca et al. (2018) and in Verbeek et al. (2018). From the zooplankton data, four variables were extracted to be used in further analyses: zooplankton total abundance, *Ceriodaphnia* abundances, *Daphnia* abundances, and copepod abundances (cyclopoids).

Bacterioplankton carbon utilization profiles using Biolog EcoPlates™

To estimate bacterioplankton carbon utilization, we used Biolog EcoPlates™ (USA). Each plate contains 31 carbon sources of different types (amino acids, carbohydrates, carboxylic acids, amines and polymers) and a blank. A complete list of the carbon sources can be found in Fig. S2 (Preston-Mafham et al., 2002). Each carbon source and the blank have three replicates, totalizing 96 wells per plate. Metabolic activity reduces a tetrazolium dye present in each well producing a specific color which is measurable as absorbance at 590 nm. 150 µL of a non-filtered, pre-sieved (200 µm mesh size) water sample was inoculated in each well with one plate per sampled pond. All incubations were done within 6 hours after the collection of the samples. From the 73 sampled ponds, we could not estimate the carbon utilization for 11 ponds, so further analyses were done based on 62 ponds. All plates were incubated for seven days in the dark at 25 °C and measured spectrophotometrically in intervals of 12, 24, 48, 72, 96, 120, 144 and 168 hours using a TECAN microplate reader. Optical densities (ODs) were corrected by subtraction of the control well OD, and negative OD values were replaced by zero. We computed the area under the curve (AUC) for each carbon source along the seven days period using the R-package *opm* (Vaas et al., 2013). AUC integrates the variation over time in metabolic activity, providing a more continuous measurement of carbon utilization (Preston-Mafham et al., 2002). To correct for differences in inoculated bacterioplankton densities, AUC values were standardized by bacterioplankton densities estimated by flow cytometry using an Attune® Acoustic Focussing Cytometer (Life Technologies, Gent, Belgium) following the same settings as described in Souffreau et al. (2014). Following the calculation of the AUC values, other curve parameters were estimated using the same *opm* package: lag phase (λ), steepness of the slope (µ) and maximum curve height (A).

Characterization of the bacterioplankton taxonomic community composition

For DNA extractions we used the UltraClean™ Soil DNA isolation kit (MoBio Laboratories Inc., Carlsbad, CA, USA). DNA was amplified by PCR using the 16S rRNA universal eubacterial

primers E338F (5'-ACTCCTACGGGAGGCAGCAGT-3') and E797R (5'-GGGTATCTAATCCTG-3') (amplicon length of 459-bp covering the variable V3 region of the 16S rRNA). The PCR mixtures were prepared with 2.5 µl 10 x PCR buffer (Eurogentec, Seraing, Belgium), 1 µl $MgCl_2$ (50 mM; Eurogentec), 2.5 µl dNTPs (2 mM; Thermo Scientific, MA, USA), 1.0 µl of each primer (20 pmol µl-1 each; Eurogentec), 0.2 µl Silverstar Taq DNA polymerase (5.0 U µl-1; Eurogentec) and 1 µl standardized template DNA (10.74 ng µl-1) in a total reaction volume of 25 µL. PCR was performed with the following settings: initial denaturation at 94 ºC for 2 min, 30 cycles of 94 ºC for 1 min, 55 ºC for 1 min and 72 ºC for 1 min, and a final extension for 5 min at 72 ºC. The quality of the PCR products was evaluated using gel electrophoresis, bands were then excised, and purification was performed using a Qiaquick Gel extraction kit (Qiagen, Venlo, The Netherlands). To quantify PCR products, we used the PicoGreen dsDNA Assay Kit (Life Technologies - USA) using manufacturer's instructions. Sequencing was performed at the Genomics Core of KU Leuven (Belgium) using a Roche 454 GS FLX+ (Switzerland).

16S rRNA pyrosequences were then processed using mothur (version 1.30.2, USA) starting with the flowgram data. The sequences contained in the flowgrams (SFF file) were sorted into groups according to the barcode sequences. For OTU (operational taxonomic unit) clustering, a 97% identity cutoff threshold was used following the standard pipeline of mothur (Schloss et al., 2009). From the 73 samples, a total of 160,430 good-quality sequences were recovered. Global singletons were removed, and samples were rarefied to the same number of sequences (1000 sequences) to compare relative differences between samples. Lundin et al. (2012) have shown that 1000 sequences per sample are a valid number to assess bacterial community variation. It remains that the number of sequences is too low to properly assess the tail of rare species in a reliable way. So, our sequencing depth is adequate for assessing the responses of the more abundant species making up the bulk of the communities, but do not inform on the dynamics of the rare fraction.

Bacterioplankton functional group annotations

In order to estimate the relative importance of assembly processes on traits with broader functional spectrum, we used FAPROTAX (Functional Annotation of Prokaryotic Taxa) (Louca et al., 2016). FAPROTAX maps prokaryotic taxa to metabolic functions using an extensive database, with the advantage to use the current literature on cultured strains. FAPROTAX provides a functional classification based on different traits and it can be used as an

exploratory tool to assess possible relations between taxonomic and functional information. It is important to note that FAPROTAX provides a picture of potential functions of communities, rather than a realized functional value (e.g., bacterial respiration rates). The python script classifies OTUs in functional groups using a database that covers different functional annotations (7600 associations) over 4600 taxa (prokaryotic – bacteria and archaea). From the 3167 OTUs in our dataset, 866 OTUs (27.35%) could be assigned to at least one functional group. From the in total 73,000 reads of the rarified dataset, 31,427 were assigned to these 866 OTUs (43.05%). The OTUs were assigned to 42 out of 90 functional groups. We focused on functions of metabolic pathways, so five functions related to biotic interactions (e.g., animal parasites and symbionts) or to human pathogens were excluded from the analyses. A complete list of the 37 functional groups used in this study can be found in Fig. S3.

Data analyses

All data analyses were performed in R (version 3.6.1) (R Core Team, 2018). First, we used Pearson's correlation in order to assess the relationship between taxonomic and functional diversity. Taxonomic diversity was estimated by number of OTUs (richness) and Shannon's diversity. Functional diversity was estimated by: (1) the number of different functional groups in the FAPROTAX classification, (2) FAPROTAX diversity (calculated using Shannon's diversity index), and (3) diversity of carbon source utilization using Shannon's diversity index, with AUC as measure of consumption. Number of the different carbon sources consumed was not used as measure of functional diversity, because relatively little variation was detected, with almost all the carbon sources being consumed to some degree by all communities (Fig. S2).

We used a variation partitioning approach to identify which factors (environment – E and/or space – S) significantly explained functional (functional group and carbon sources utilization) and taxonomic variation using R-package *vegan* (Oksanen, 2015). For this, a combination of redundancy analysis (RDA) and forward selection was used. Previous to the analyses, correlations (Spearman's rank) were tested across all environmental variables. Environmental variables were log or logit (for percentages) transformed to achieve normality when necessary. For the redundancy analysis, all environmental variables were standardized (z-scores). To calculate spatial variables, we generated Moran's eigenvector map (MEM) from the measured latitude and longitude positions of the ponds (Beisner et al., 2006; Dray et al.,

2006; Yeh et al., 2015). RDA analysis was first performed on all variables (global test) for each set (E and S) separately using a permutational test of significance (1000 permutations) and only if the global analysis was significant ($p < 0.05$), forward selection was applied. To avoid collinearity all variables were tested for variance inflation factors (VIFs < 10) before forward selection.

RDA was performed for the different taxonomic and functional datasets separately. For the taxonomic abundance-based dataset (OTUs), we used Hellinger transformation to reduce the impact of zeros previous to the analysis (Legendre & Gallagher, 2001). The FAPROTAX dataset (functional group annotations) was transformed to relative abundances of the functional records based on the 866 OTUs' relative abundance data before the RDA analysis. For the carbon utilization (Biolog EcoPlates™), area under the curve values were standardized by log-transformed bacterial densities. Since functional and taxonomic datasets had a considerable difference in the number of response variables (two orders of magnitude), we also performed a distance-based RDA (db-RDA). Distance based RDA was carried out using PCoA eigenvectors as response variables for both functional and taxonomic datasets. Bray-Curtis dissimilarity was used for the PCoA calculation followed by *lingoes* correction when negative values were generated by the dissimilarity matrices (Borcard et al., 2011). The PCoA eigenvectors used as response variables were selected based on values of F-statistic and adjusted R^2 values of the different combinations of response variables and E or S (Anderson & Willis, 2003; Duarte et al., 2012; Gianuca et al., 2018).

Variation partition analysis was performed using the set of variables (E and S) selected on the RDA and db-RDA analyses. Using a partial RDA analysis, variation partition provides the relative contribution of the conditional effect of environment corrected for space and vice-versa (Peres-Neto et al., 2006). Additionally, a shared component between E and S can be quantified. Conditional effects were tested by permutation (1000 permutation) using the *anova.cca* function of the *vegan* package.

To better interpret the spatially structured variation, we first performed RDA on the selected set of MEMs and the community structure data and extracted the significant canonical axes using the *anova.cca* function. This way, we generated axes that capture the variation in community structure explained by space. Then, the fitted scores of the generated canonical axes were plotted on their respective geographic locations in a map of sites to visualize the spatial patterns in community variation. Additionally, we tested if the response

variation (taxonomic and functional) explained by space was related to measured spatially-structured environmental data. To do so, the fitted site scores of the selected canonical axes were regressed with all measured environmental variables (Borcard et al., 2011).

The influence of the selected environmental variables on the taxonomic structure was visualized and analysed using bar plots and Pearson's correlations at phylum and class level. For the functional structure, Pearson's correlation was used to assess the relation between the selected environmental variables and each specific metabolic function. Since the RDA analysis of the carbon consumption profiles did not select any environmental variable, this analysis was not performed for this dataset.

Other graphical figures (e.g. barplots) were generated using different R-packages: *ggplot2* (Wickham, 2009), *GGally* (Schloerke et al., 2018), *ggfortify* (Tang et al., 2016), *phyloseq* (McMurdie & Holmes, 2013), *RColorBrewer* (Neuwirth, 2014), and *viridis* (Garnier, 2017). A map of Belgium (Fig. S1) was produced using packages *maps* (Becker et al., 2018) and *BelgiumMaps.StatBel* (https://github.com/bnosac/BelgiumMaps.StatBel).

4.4 Results

Taxonomic and functional bacterioplankton community structure

Twenty different phyla were found over the 73 ponds. Proteobacteria (total relative abundance: 46.37%) was the most abundant phylum followed by Actinobacteria (24.08%), Bacteroidetes (14.48%) and Firmicutes (1.76%) (Fig. S4). 11.59% of the taxa could not be classified at phylum level. At class level, 39 different classes were observed. Among the 39 classes, Betaproteobacteria (28.19%) was the most abundant class followed by Gammaproteobacteria (7.80%), Alphaproteobacteria (7.72%), Sphingobacteria (7.12%) and Flavobacteria (4.34%) (Fig. S5). 15.1% of the OTUs could not be assigned to any taxon at class level.

For the carbon consumption profiles measured by Biolog EcoPlates^{TM}, all 31 tested carbon sources were metabolized to some degree by all bacterioplankton communities (Fig. S2). L-Asparagine was the most metabolized carbon source across all samples in terms of area under the curve (AUC) values, followed by N-Acetyl-D-Glucosamine, L-Arginine and D-Cellobiose (Fig. S2). Curve parameters estimated for each carbon source across samples are summarized as follows (given as optical density at 590 nm): area under the curve (AUC) (mean

= 133.48 OD_{590} ± 28.32 SD), maximum curve height (A) (1.62 OD_{590} ± 0.27), steepness of the slope (μ) (0.032 OD_{590} ± 0.004) and lag phase (λ) (40.73 OD_{590} ±10.89).

Functional groups assigned by FAPROTAX were distributed over 37 different metabolic groups (Fig. S3). Most common functional groups across samples were mainly represented by chemoheterotrophy (total relative abundance: 24.22%) followed by aerobic chemoheterotrophy (15.53%), methylotrophy (7.11%), hydrocarbon degradation (4.77%) and methanotrophy (4.76%).

Relationship between taxonomic and functional diversity

To investigate the relationships between taxonomic and functional diversity, we used Pearson's correlation. The number of functions assigned with FAPROTAX was significantly positively related to the number of OTUs (r = 0.50, $p < 0.001$) (Fig. 1a). Functional group diversity was also positively related to taxonomic diversity (Shannon's diversity of OTUs), but showed a weaker relation (r = 0.27, $p = 0.017$) (Fig. 1b). Functional diversity of carbon utilization (also calculated as Shannon's diversity) was not significantly related to taxonomic diversity (r = 0.05, $p = 0.66$) (Fig. 1c).

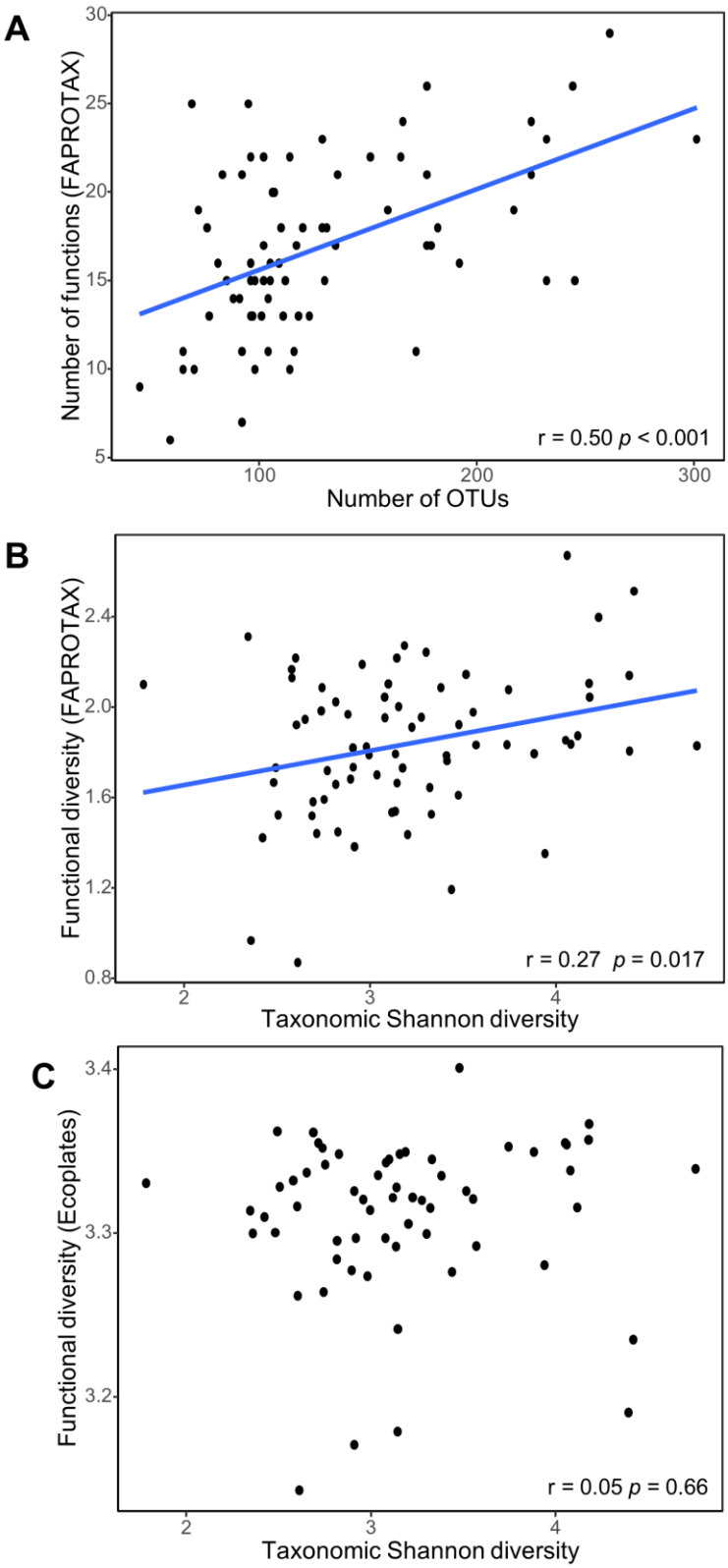

Figure 1. Correlations between: (A) number of functions (FAPROTAX) and OTU richness (number of different OTUs); (B) functional diversity (calculated as Shannon diversity of the function annotations based on FAPROTAX) and taxonomic Shannon diversity (Shannon diversity of OTUs); (C) functional diversity (calculated as Shannon diversity of the consumption of the 31 carbon sources) and taxonomic Shannon diversity (Shannon diversity of OTUs). Pearson's correlation (r) and *p*-values of the correlation are given for each plot.

Relative importance of environmental and spatial drivers shaping taxonomic and functional community structure

For the Hellinger-transformed taxonomic data, environmental variables explained 11.4% (R^2adj values) (Table S1) of the variation with five variables being selected by forward selection: suspended matter, pH, available phosphorus, *Daphnia* abundances and total phosphorus. Spatial variables explained 5.5% of the variation with MEM variables ranging from broad (MEM4 and MEM5) to fine-scale spatial variables (MEM24, MEM25 and MEM38). Using the environmental and spatial variables selected by forward selection, we applied a variation partitioning approach based on partial RDA analyses to quantify the conditional effect of each component (i.e., environment corrected for space - E|S and space corrected for environment S|E). For the taxonomic structure based on Hellinger-transformed abundances, pure environment (E|S, R^2adj = 8.9%) was still more important than the spatial component (with all variables remaining significant when corrected for space), but pure space (S|E, R^2adj = 3%) also explained a significant fraction of variation; 2.6% of the variation being shared by the two components (Fig. 2a). For the distance-based taxonomic structure (represented by three PCoA eigenvectors), environment was also more important (R^2adj = 35.1%) (Table S1), with five variables being selected: suspended matter, pH, *Daphnia* abundances, available phosphorus and maximum depth. A total of, 17.8% of the community variation was explained by two broad and two fine-scale spatial variables (MEM4, MEM5, MEM24 and MEM27). Variation partitioning analysis showed that pure environment (E|S) explained 23.1% (R^2adj) of the variation (maximum depth was not significant when corrected by space), while pure space (S|E) explained 5.8%; the shared component (E ∩ S) explained 12% of the total variance (Fig. 2b).

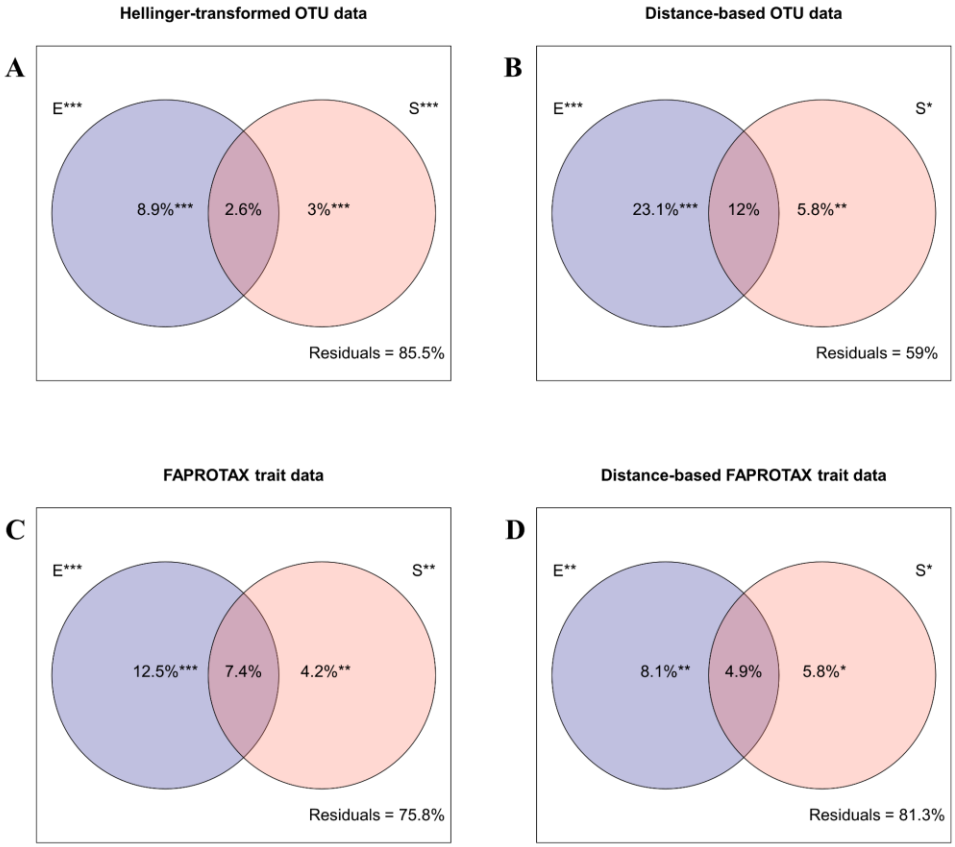

Figure 2. Venn diagrams representing the results of variation partitioning analyses on taxonomic and functional bacterioplankton community composition. Variation explained (adjusted R^2 - ***$p < 0.001$, **$p < 0.01$, *$p < 0.05$) by environmental (E) and spatial (S) variables is shown for (A) Hellinger-transformed OTU data; (B) Bray-Curtis distance-based OTU data; (C) relative functional group abundances (FAPROTAX trait data) and (D) Bray-Curtis distance-based functional group (distance-based FAPROTAX trait data). Residuals are the unexplained variance for each dataset.

Relative functional group abundances based on FAPROTAX annotations were also driven mostly by environment, as 19.9% of the total variation was explained by dissolved oxygen, total and available phosphorus (Table S1). 11.6% of total variance was explained by two broad-scale spatial variables (MEM5 and MEM6). When corrected for space, pure environment still explained most part of the variation (12.5%) (Fig. 2c), with the shared effect (7.4%) explaining more than pure space (4.2%). Among the environmental variables, available phosphorus did not remain significant when corrected for space. When analyzing the functional group annotations dataset using the distance-based approach (represented by two PCoA eigenvectors), 12.9% of variation was explained by two environmental variables: suspended matter and dissolved oxygen (Table S1). Spatial signal explained 10.6% of the total variation (one broad and one fine-scale spatial variable: MEM5 and MEM31). For the distance-

based FAPROTAX dataset, pure environment explained 8.1% of the variation, with pure space (5.8%) and the shared effect (4.9%) explaining smaller portions of variation (Fig. 2d).

Variation in carbon utilization profiles (measured by Biolog EcoPlates™) could not be explained by any of the environmental or spatial variables as the RDA global test was not significant for any component (Table S1).

The spatial (S) and the shared spatial and environmental component (E ∩ S) were further investigated using canonical axes generated by RDA on the selected set of MEMs and the community structure data to capture the spatially structured variation in community structure. This spatially structured community variation (fitted scores of the generated canonical axes) was plotted on their respective geographic locations, and we tested for relations among the fitted scores of these canonical axes and the environmental variables using linear regression. For the Hellinger-transformed OTU data, two canonical axes were generated for the spatially structured community variation (RDA1-tax and RDA2-tax) (Fig. S6a and S6b). RDA1-tax was significantly related to total zooplankton abundance (t-value = 2.208, p = 0.047) and to the abundances of *Ceriodaphnia* (t-value = -2.307, p = 0.024). RDA2-tax was significantly related to transparency (t-value = 2.449, p = 0.017), pH (t-value = -2.582, p = 0.012) and suspended matter (t-value = 3.376, p = 0.001). For the distance-based OTU data, three canonical axes (RDA1-taxdist, RDA2-taxdist and RDA3-taxdist) were generated from the combinations of the selected MEMs (Fig. S6c, S6d and S6e). RDA1-taxdist was significantly related to the concentrations of available phosphorus (t-value = 2.637, p = 0.010) and to the abundances of *Ceriodaphnia* (t-value = 2.211, p = 0.031). RDA 2-taxdist was not significantly related to any environmental variable and RDA3-taxdist was significantly related to pH (t-value = -2.263, p = 0.027) and to suspended matter values (t-value = 2.912, p = 0.005). Regarding the FAPROTAX trait data, the combination of the selected MEMs and community structure generated one significant canonical axis (RDA1-faprotax) (Fig. S7a), which was significantly related to transparency (t-value = -2.073, p = 0.042) and suspended matter (t-value = -2.901, p = 0.005). Distance-based FAPROTAX trait data also had one (RDA1-faprotaxdist) significant canonical axis generated by the combination of the selected MEMs (Fig. S7b). This canonical axis was significantly related to different environmental variables: transparency (t-value = -2.491, p = 0.015), pH (t-value = 2.832, p = 0.006), suspended matter (t-value = -2.932, p = 0.004) and abundances of *Daphnia* (t-value = 2.151, p = 0.035). These results indicate that

there is a spatial structure for these specific selected environmental variables, influencing community structure.

To capture how the different communities (OTU data) and functional groups (FAPROTAX data) are distributed over the selected environmental variables, we used RDA plots (Fig. 3). For the Hellinger-transformed OTU data, suspended matter and total phosphorus formed a long environmental gradient most likely related to a gradient in eutrophication (Fig. 3a). Environmental variables selected for the distance-based OTU data also separated communities along an axis of suspended matter concentration and abundance of *Daphnia* (Fig. S8a). Relative functional group abundances (FAPROTAX) were driven by a gradient of total phosphorus and dissolved oxygen (Fig. 3b). The most abundant functional groups (chemoheterotrophy, aerobic chemoheterotrophy, methylotrophy, hydrocarbon degradation and methanotrophy) were related to these two selected environmental variables. However, a large number of functions did not show a clear relation to the selected environmental variables (Fig. 3b). Distance-based FAPROTAX trait data followed the same pattern with communities separated along an axis of suspended matter concentration and dissolved oxygen (Fig. S8b).

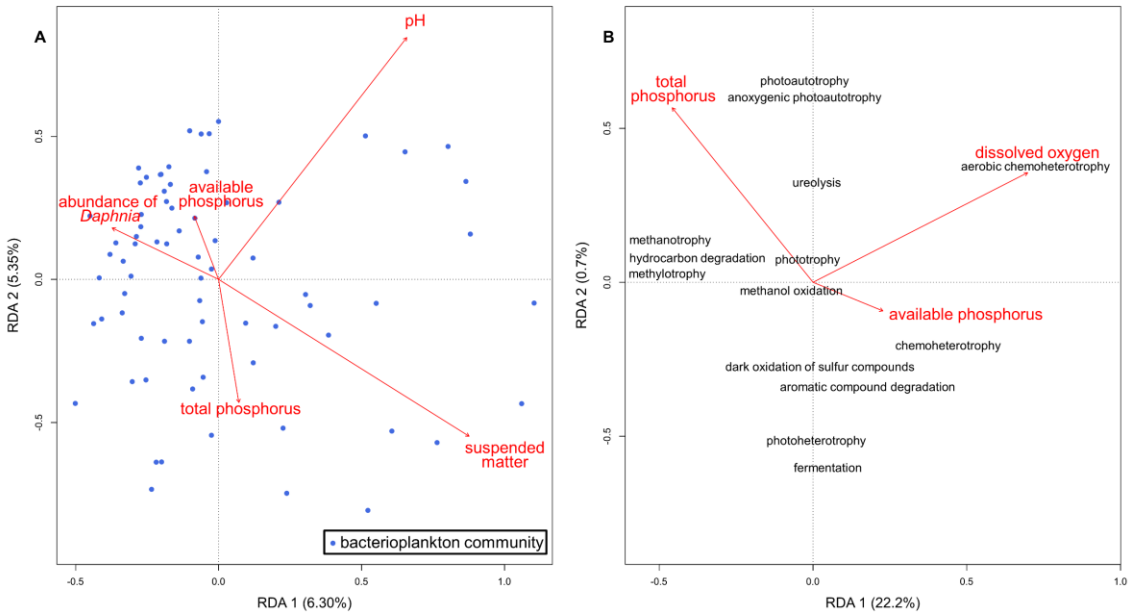

Figure 3. Redundancy analysis plots of the environmental variables selected by forward selection for: (A) taxonomic Hellinger-transformed abundance based (3167 OTUs) – each dot depicts a bacterioplankton community; (B) relative functional group abundances (37 FAPROTAX functions) – functional groups with less than 1% of relative abundance were omitted due to graphical constraints.

Associations of taxa and functional groups to environmental variables

We used the environmental variables selected by the redundancy analyses to address specific changes at phylum and class level for the taxonomic structure. Bar plots were used to place each of the 73 communities at phylum (Fig. S4) and class (Fig. S5) level ordered by the values of suspended matter (explaining the highest amount of variation of the taxonomic structure). At phylum level, several selected environmental variables significantly correlated to the main phylum taxa: Proteobacteria was positively correlated to the suspended matter values (r = 0.32, p = 0.005) while the relative abundance of Actinobacteria presented the opposite pattern (r = -0.37, p = 0.001). Proteobacteria (r = -0.30, p = 0.008) and Actinobacteria (r = 0.44, p < 0.001) were also significantly correlated with the abundance of *Daphnia*. Again, relative abundance of Proteobacteria was correlated to the total phosphorus concentration (r = 0.23, p = 0.04) and Bacteroidetes was positively correlated to pH (r = 0.46, p < 0.001). Among the most abundant classes, only the relative abundance of Sphingobacteria showed a significant correlation to the values of suspended matter (r = -0.33, p = 0.004). pH was correlated with different classes – Gammaproteobacteria (r = -0.28, p = 0.015), Alphaproteobacteria (r = 0.26, p = 0.021) and Sphingobacteria (r = 0.44, p < 0.001). Sphingobacteria was also negatively correlated with total phosphorus (r = -0.24, p = 0.039). While available phosphorus was significantly correlated with Alphaproteobacteria (r = -0.24, p = 0.033) and Flavobacteria (r = 0.26, p = 0.022). Lastly, Alphaproteobacteria was negatively correlated with abundance of *Daphnia* (r = -0.30, p = 0.008).

We used the four environmental variables (dissolved oxygen, suspended matter, total and available phosphorus) selected by RDA models to explain specific correlations with the relative abundance of each metabolic function. Significant correlations were related to the most common metabolic functions (Fig. 4). Aerobic chemoheterotrophy, chemoheterotrophy and aromatic degradation were positively correlated to dissolved oxygen and negatively correlated to total phosphorus, while methanotrophy, methylotrophy and hydrocarbon degradation were negatively correlated to dissolved oxygen and positively correlated to total

phosphorus concentration. A complete map of correlations of the metabolic functions and environmental variables is shown in Fig. 4.

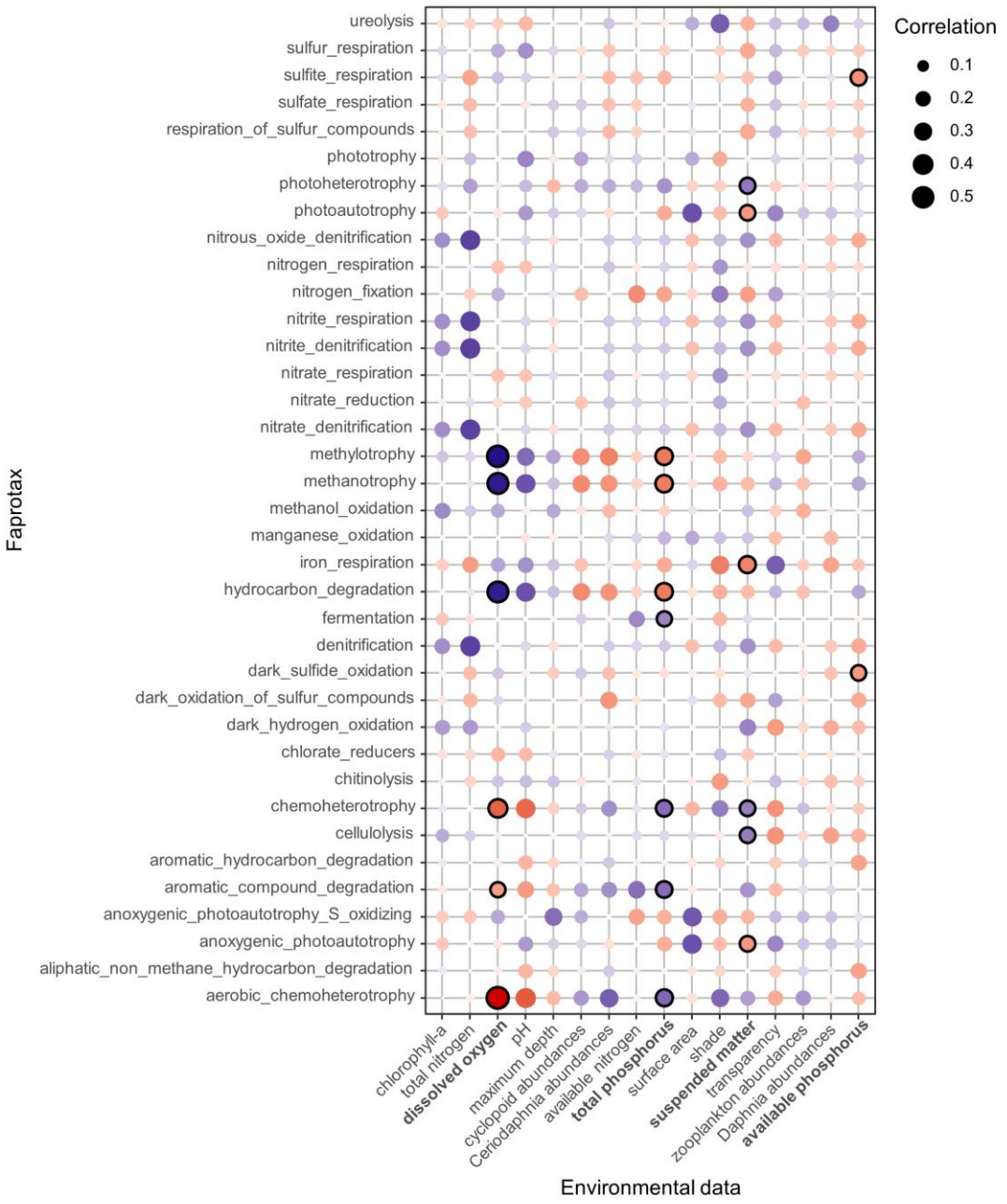

Figure 4. Pearson's correlations of functional groups (assigned by FAPROTAX) and environmental variables. Colors depict negative (blue) to positive (red) correlations. Dark circles represent significant correlations (p < 0.05).

4.5 Discussion

We assessed the diversity and assembly mechanisms of bacterioplankton communities from shallow ponds in Belgium at the taxonomic and functional level, using two trait-based estimations (functional group annotations and carbon utilization profiles) to assess functional composition. Only for the functional group annotations was functional diversity positively related to taxonomic diversity. The community structure at both the taxonomic level and the level of functional group annotations was significantly associated with environmental conditions, yet with a different set of environmental variables explaining functional and taxonomic community variation. Some environmental variables were directly associated to changes in the relative abundances of specific functional and taxonomic groups. Spatial variables also played a significant role for both the functional group structure and for taxonomic structure, but their relative importance was lower compared to that of environmental variables. Community structure based on carbon utilization profiles could not be explained by environment or space.

Functional group diversity was significantly correlated to taxonomic diversity (number of OTUs and Shannon diversity), both in terms of number of functional groups as well as when integrating relative abundances of the functional groups (Shannon index). This is similar to the pattern reported for marine bacteria and archaea across the global ocean (Louca et al., 2016). An increasing number of studies suggest that microbial diversity enhances ecosystem functioning (Horner-Devine et al., 2003; Venail & Vives, 2013). The observed relationship between microbial diversity and functional group annotations in shallow ponds by us is in line with these results. Conversely, functional diversity estimated based on the carbon utilization profiles did not show a significant correlation to taxon diversity. The lack of relation between taxonomic and carbon utilization diversity reflects functional redundancy (Comte & del Giorgio, 2010; Konopka, 2009).

The community structure at both the taxonomic level and functional group annotations level was mainly explained by local environmental pond conditions, yet by a different set of environmental variables. Community structure based on functional group annotations was mainly driven by environmental variables related to gradients of nutrients (total and available phosphorus) and dissolved oxygen (Fig. 2). Some of the more abundant functional groups (methylotrophy and methanotrophy) were associated with high nutrient

conditions (Fig. 3b and Fig. 4). Taxonomic structure was driven by the same gradient of nutrients, but pH and abundance of *Daphnia* also explained part of the variation (Fig. 3a). pH is well known to drive bacterial community composition in freshwater systems and soil (Xu et al., 2014; Ren et al., 2015; Hanashiro et al., 2019), and the cladoceran zooplankton genus *Daphnia* has been described as an important biotic factor related to bacterioplankton taxa (Souffreau et al., 2015; Hanashiro et al., 2019). *Daphnia* are known to regulate bacterioplankton communities as grazers (Birtel & Matthews, 2016; Degans & De Meester, 2002; Jürgens et al., 1994; Jürgens & Jeppesen, 2000) and through their symbiotic relationships (Callens et al., 2016).

Compared to environmental variables, spatial variables explained a lower but still significant portion of community variation at both taxonomic and functional level, and similar spatial variables were selected for the two community indicators. A spatial community structure can indicate that dispersal limitation or mass effects are structuring communities (Cottenie, 2005). Yet, a significant spatial signal could also be related to an unmeasured spatially-structured environmental variable, or to local biotic interactions through other trophic levels that are spatially structured (Verreydt et al., 2012). The absence of a distance-decay relationship between community variation and distance (negative exponential model - pseudo R^2 = 0.0002, p = 0.377) suggests that dispersal limitation might not play a dominant role in the studied communities. The shared effect of environment and space is related to measured environmental variables that are spatially structured (Bell, 2010; Souffreau et al., 2015). In this study, three abiotic variables (transparency, pH and suspended matter) and three biotic variables (total zooplankton abundance, abundances of *Ceriodaphnia* and *Daphnia*) were correlated to the spatially related community structure (taxonomic and functional). This latter indicates a spatial structure for the zooplankton communities, which consequently could affect the bacterioplankton community structure. Zooplankton communities are known to affect freshwater bacteria both directly through grazing and indirectly through increase in substrate availability by algae consumption (Šimek et al., 2014). Inspection of the patterns of Fig. S6 suggests that general soil characteristics might affect bacterioplankton communities, as the patterns show a distinction between the region of Western and Eastern Flanders (to the West; dominated by sandy loam and sandy soils), the Campine area (to the right) that is more dominated by sandy soils, and Central Belgium and

the Walloon area, which have variable soil characteristics but differ from both other regions (Meersmans et al., 2011).

The dominant bacterial taxa were probably driven by axes of variation related to three main factors: pH, productivity and zooplankton biomass and community composition (Fig. 3a). pH has been described as an important driver for microbial communities, responsible for species turnover in different conditions, from continuous gradients (Hanashiro et al., 2019) to more extreme conditions (Ren et al., 2015). In our study, pH conditions ranged from 5.69 to 10.14, and the phyla Bacteroidetes and the classes Alphaproteobacteria and Sphingobacteria were significantly positively related to the pH gradient. In terms of the productivity gradient, part of the bacterioplankton communities were driven by increase in values of suspended matter and total phosphorus concentrations (Fig. 3a). More productive aquatic environments could allow more species to flourish and reach higher densities (Smith, 2007). Moreover, some specific taxa can profit from these more productive conditions and become dominant, that way driving community patterns. In our study, the phyla Bacteroidetes and the classes Alphaproteobacteria and Sphingobacteria were indeed positively associated with nutrient conditions. Interestingly, the abundance of *Daphnia*, a known grazer of bacteria, was in our study significantly correlated with specific bacterioplankton phyla (Actinobacteria and Proteobacteria) (Jürgens et al., 1994). This could explain the contrasting patterns found for different bacterioplankton phyla: *Daphnia* abundances were positively related to the relative abundances of Actinobacteria, but negatively to Proteobacteria abundances.

Also certain functional groups were related to specific environmental conditions. Exclusive aerobic chemoheterotrophs were distributed following increasing concentrations of dissolved oxygen, while chemoheterotrophs were spread across most of the shallow ponds (Fig. S3). Chemoheterotrophs convert chemical elements from organic to inorganic elements and cover a large part of bacterioplankton taxa in water bodies (Konopka, 2009). In our study, taxonomic structure of bacterioplankton communities and their functional group annotations were driven by a different set of environmental variables (Table S1). Using genomic traits of ocean microbial communities, Barberán et al. (2012) also found a mismatch between what drivers were important for the taxonomical and functional structure patterns. This could reinforce the idea of a high level of functional redundancy even among taxonomically distinct communities. Moreover, this highlights the importance of including the functional approach

to more fully understand the patterns related to environmental filtering on the assembly of microbial communities.

In our study, carbon utilization profiles (Biolog EcoPlates™) variation could not be explained by the measured environmental and spatial variables. While most of the studied carbon sources were consumed in similar proportions across samples, the relative abundances of functional group annotations showed a more heterogeneous distribution along the gradient of dissolved oxygen. The lack of an association of carbon source utilization and environmental variables may reflect that we did not measure carbon source profiles in all ponds; in a study that included dissolved organic matter (DOM), a relationship between a gradient of quality of DOM and carbon use profiles was observed (Ruiz-González et al., 2015). It might also be linked to the fact that carbon utilization relies on relatively simple genetic pathways (i.e. encoded by few genes) which are likely shared by most bacterioplankton taxa (Martiny et al., 2013). Bacterial functional traits are known to be conserved at different phylogenetic depths (Martiny et al., 2009; Martiny et al., 2015). pH resistance, for example, is shared among taxa at relatively deep phylogenetic relatedness. Generally, complex traits encoded by a high number of genes, such as methanogenesis, are related at deep phylogenetic levels. More simple traits, such as the consumption of simple carbon sources tend to be phylogenetically overdispersed (Martiny et al., 2013). This is thought to lead to a weaker connection between carbon utilization profiles and any environmental condition (Bissett et al., 2010; Lima et al., 2016). Overall, our results suggest that trait-based approaches might add an important layer of information to understand how communities respond to environmental changes (Menezes et al., 2010; Tiquia, 2010) and that biogeographical studies might improve their predictability using traits rather than looking only to the taxonomic structure.

We are well aware of the limitations associated with the functional approaches used in our study. Carbon utilization profiles generated by Biolog EcoPlates™ have been criticized due to the fact that the consumption of one single carbon source would lead to a dominance and growth of specific taxa and this would no longer represent the functional potential of the original microbial community (Preston-Mafham et al., 2002). We tried to overcome this problem using the area under the curve (AUC) analysis, which takes into consideration the whole period of incubation, thus compensating for possible shifts in the dominance of the species consuming the substrate. In terms of the functional profiles based on the FAPROTAX

approach, a considerable number of OTUs were not assigned to any metabolic function. However, it is important to note that in terms of sequence reads, almost half of the sequences were assigned to at least one metabolic function. Most abundant OTUs are represented by a functional group annotation in our analyses, while rare taxa, which is the majority of the OTUs not assigned to a metabolic function, were also not considered in the taxonomic analyses. We understand that this assignment issue reduces the estimation of metabolic diversity in our dataset, however, our study provides an exploratory picture of how these functional traits are been selected in these environments. Moreover, it has been shown that dominant taxa are often responsible for the ecological functions provided by a community (Yang, 2021). Recently though, the importance of the inclusion of rare taxa in microbial assembly studies has been reported (Liao et al., 2017). It has been shown that in variation partitioning, most variation is driven by the most abundant taxa/functions. Thus, both the taxonomic and functional analyses focus on the more abundant taxa. As such, the comparison between the functional and taxonomical profiles are not strongly impacted by the exclusion of the rare taxa/functions. We are, however, fully aware of the fact that our analyses only provide information on the patterns driven by the more abundant taxa.

Overall, our study on shallow freshwater ponds in Belgium shows that functional group annotations and taxonomic structure of bacterioplankton communities were driven mainly by environment, with different sets of variables being important for each. Spatial variables also explained part of the variation for functional groups and taxonomy and were potentially linked to regional environmental differences. A large part of variation remained unexplained. Part of this unexplained variation might be related to stochastic processes (e.g. drift) (Zhou & Ning, 2017; Langenheder & Lindström, 2019). Our results combining taxonomic and functional bacterioplankton indices improved our interpretation on how taxonomic identity and functional profile respond differently to local environmental conditions, thus contributing to a better understanding on how these organisms trace and adapt to the local conditions.

Acknowledgments

We thank Kristien Brans, Jessie Engelen, Matthias Vanhamel, Edwin van den Berg, Andros Gianuca and Carla Denis for their contributions to the field work and laboratory analyses. We also thank the different SPEEDY and TippingPond partners for their help and their contributions to the general development of the projects.

4.6 Supplementary material for Chapter IV

Table S1. Summary of the RDA analyses showing the relative importance of the different components of explained variation (adjusted R^2) of community composition in the Hellinger-transformed abundance dataset (3167 OTUs), relative functional group abundances (37 FAPROTAX functions) and carbon utilization (31 EcoPlates™ sources). Explained variation (adjusted R^2) was also calculated for distance-based RDA for each of the three datasets. The effects of each component (E – environment and S – space) are not (marginal effect) or are corrected (conditional effect) for the other components. NS means that a certain model did not explain any variation in community composition significantly, or that the % explained variance was not significant. Environmental variables in bold remained significant after correction for effects of space.

Model	Components studied					
	Marginal		Conditional			
	E	S	E|S	S|E	Significant environmental variables	MEMs
Hellinger-transformed abundances (OTUs)	11.4%***	5.5%***	8.8%***	3%***	**Suspended matter**, **pH**, **available phosphorus**, **abundance of *Daphnia*** and **total phosphorus**	5, 24, 28, 25, 38
Distance-based (OTUs)	35.1***	17.8%*	23.1%***	5.8%**	**pH**, **suspended matter**, **abundance of *Daphnia***, **available phosphorus** and maximum depth	4, 5, 24, 27
Relative functional group abundances (FAPROTAX)	19.9%***	11.6%**	12.5%***	4.2%**	**Dissolved oxygen**, available phosphorus and **total phosphorus**	5, 6
Functional group distance-based (FAPROTAX)	12.9%**	10.6%*	8.1%**	5.8%*	**Dissolved oxygen** and **suspended matter**	5, 31
Carbon utilization (EcoPlates™)	NS	NS	-	-	-	
Carbon utilization distance-based (EcoPlates™)	NS	NS	-	-	-	

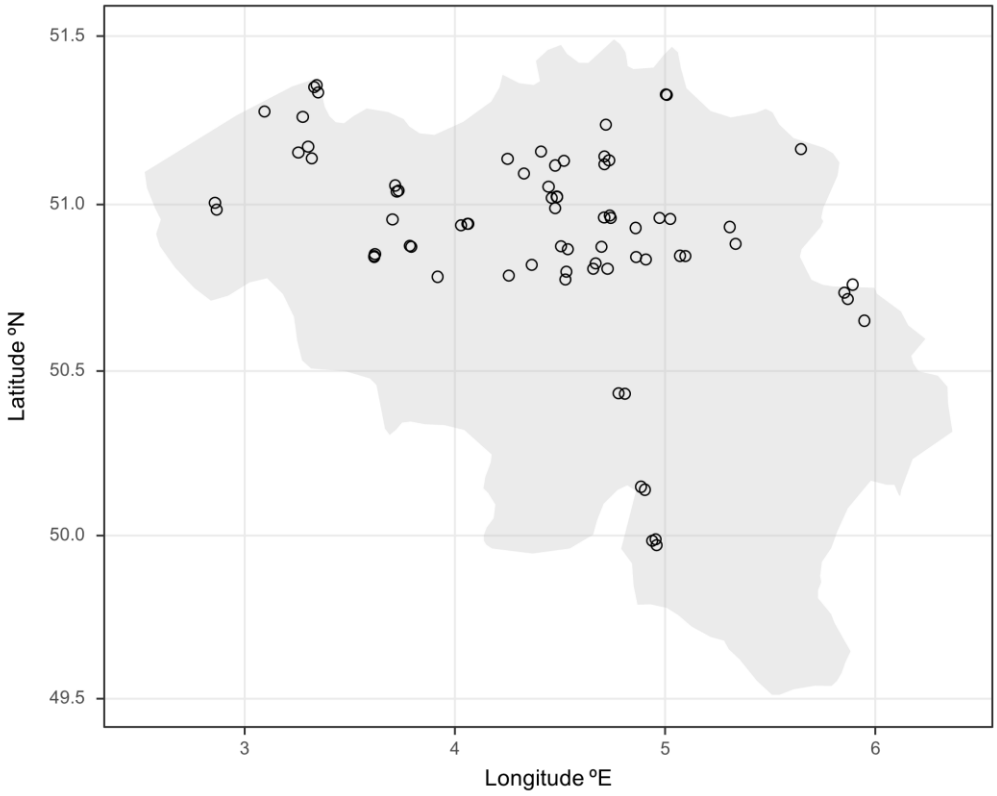

Figure S1. Map showing the location of the 73 shallow ponds sampled across Belgium.

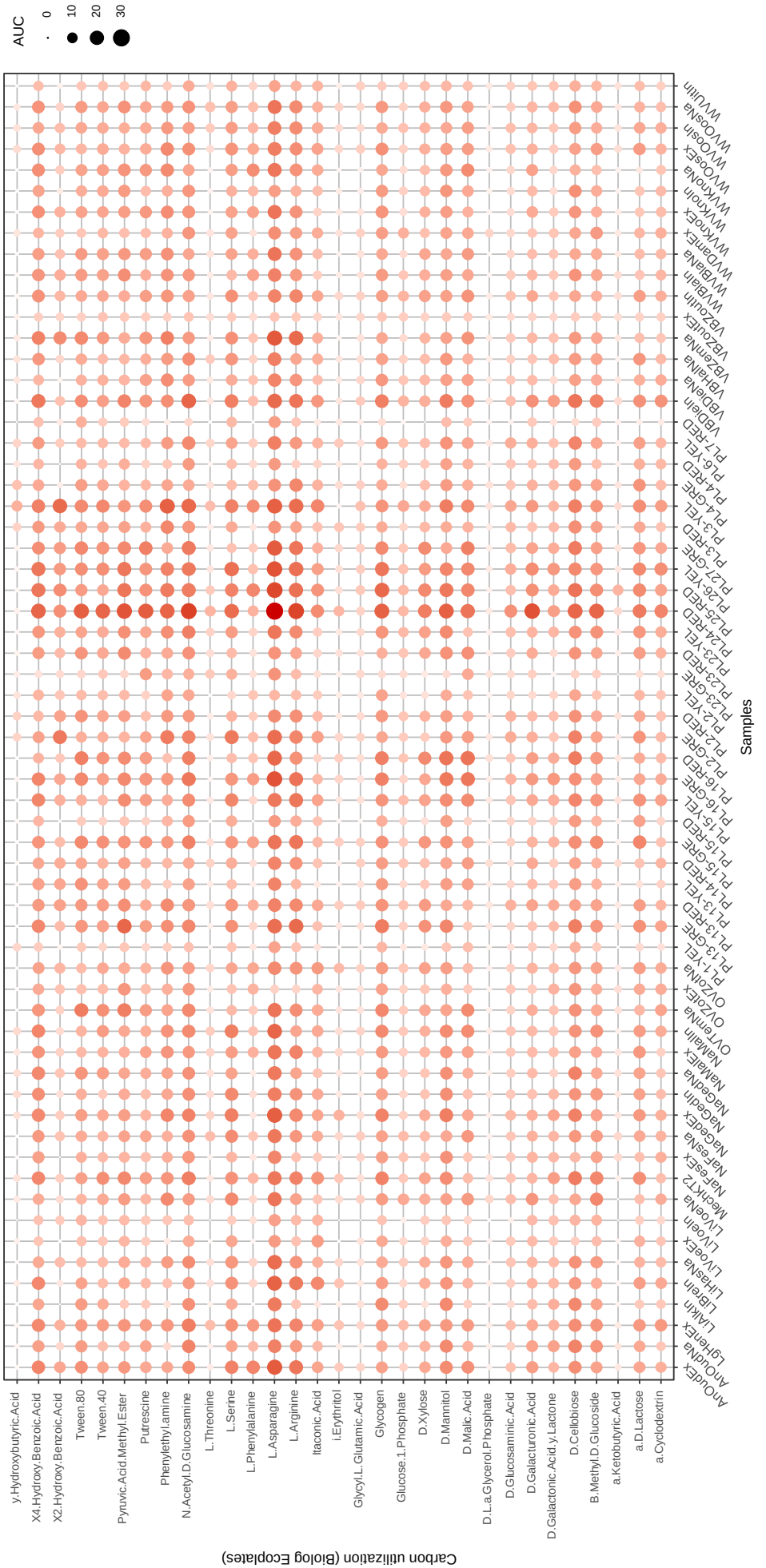

Figure S2. Carbon utilization (Biolog Ecoplates) for the 31 different carbon sources along the 62 shallow ponds sampled in Belgium. Circle size and color saturation represent the area under the curve (AUC) values for each carbon source. Samples (X axis) are ordered alphabetically.

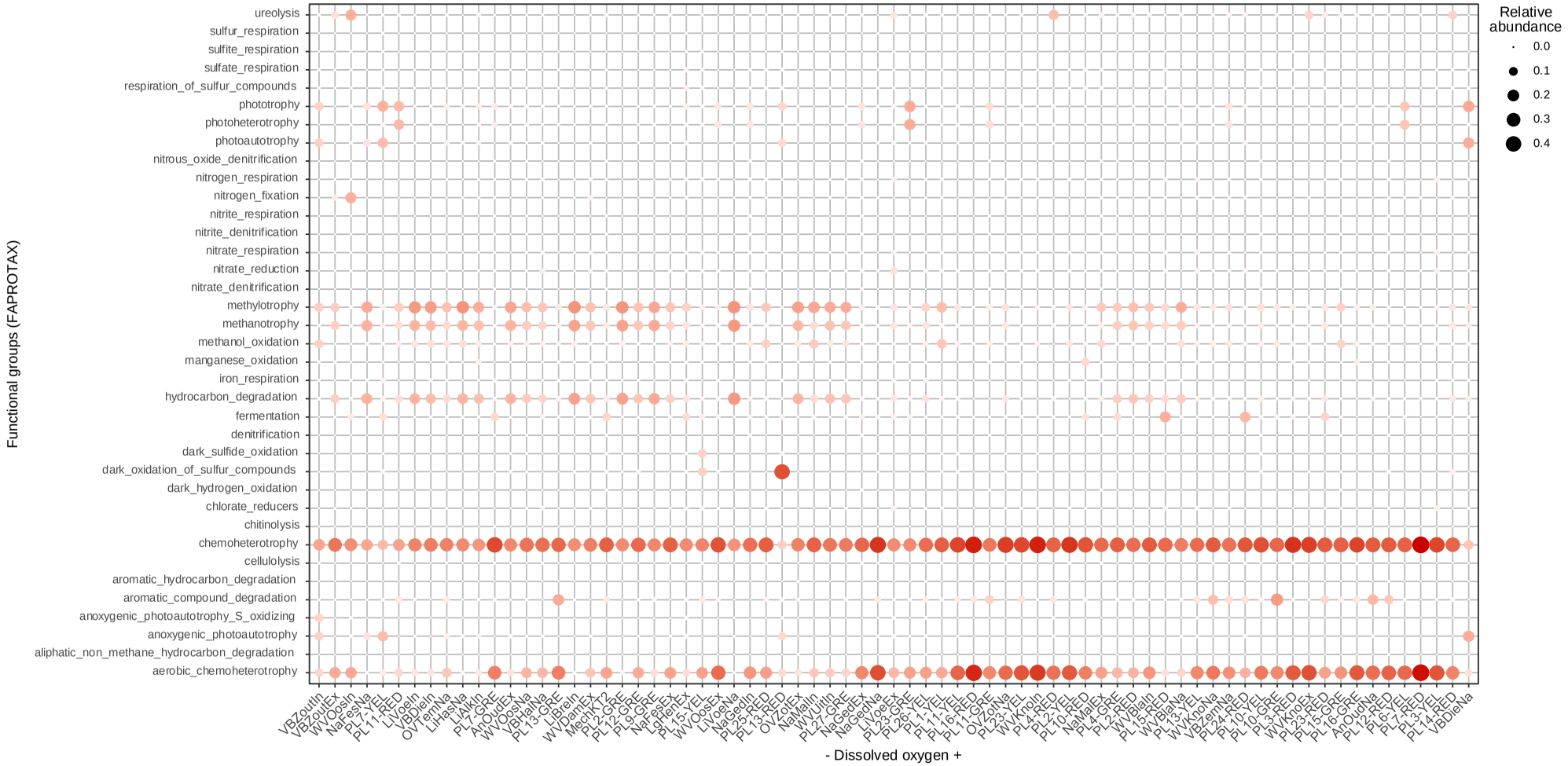

Figure S3. Relative functional group abundances (FAPROTAX) along 73 shallow ponds sampled in Belgium. Circle size and color saturation are proportional to relative record values for functional groups. Samples (X axis) are ordered by dissolved oxygen values.

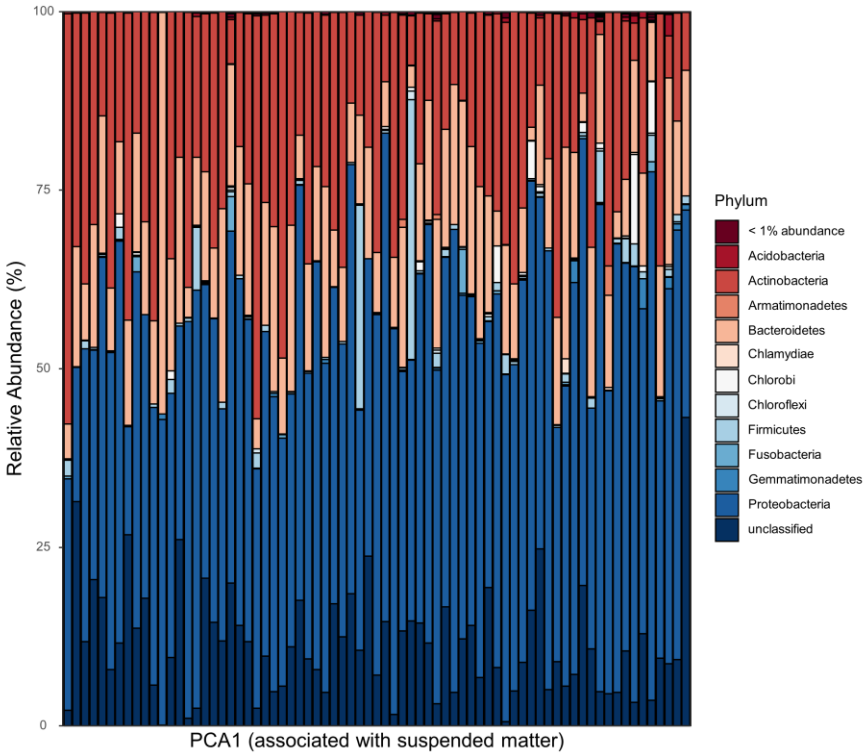

Figure S4. Relative abundance profiles at the phylum level of the 73 shallow ponds. Bars are ordered by the first axis of the principal component analysis, which was related to suspended matter concentration. Phylum taxa with relative abundances lower than 1% were grouped together.

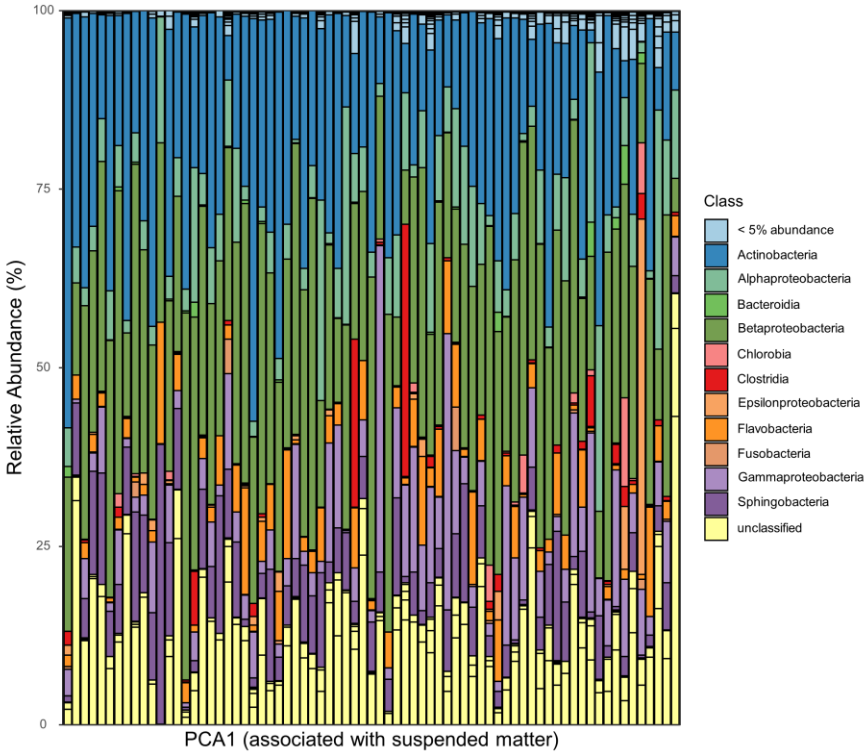

Figure S5. Relative abundance profiles at the class level of the 73 shallow ponds. Bars are ordered by the first axis of the principal component analysis, which was related to suspended matter concentration. Class taxa with relative abundances lower than 5% were grouped together.

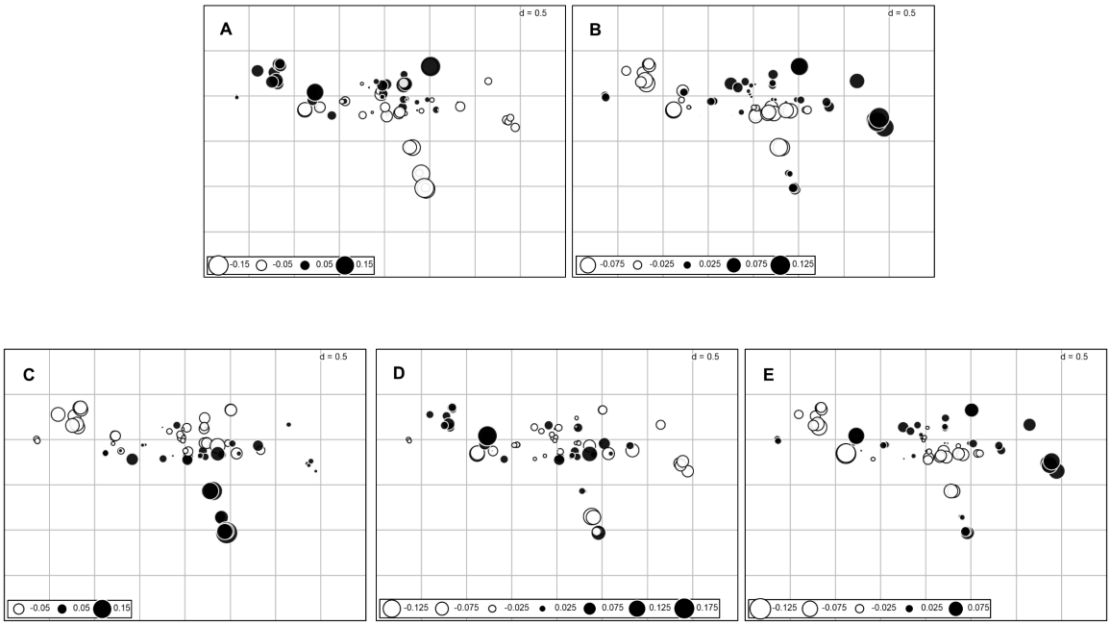

Figure S6. Maps of the fitted site scores of the canonical axes (i.e., community variation) generated for the combinations of the *a priori* selected MEMs, thus representing the variation in community structured explained by combinations of selected geographical variables: (A) and (B) – RDA1 and RDA2, respectively, for the Hellinger-transformed OTU data; (C), (D) and (E) – RDA1, RDA2 and RDA3, respectively, for the distance-based OTU data. Positive values are represented by filled circles and negative values by open circles. Their sizes illustrate their absolute values, and their positions refer to the geographical position of each sampled shallow pond.

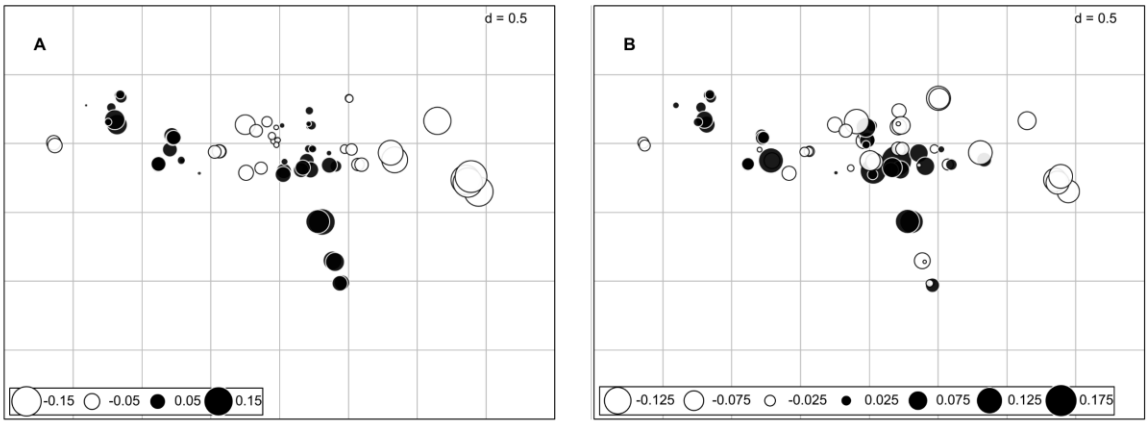

Figure S7. Maps of the fitted site scores of the canonical axes (i.e., community variation) generated for the combinations of the *a priori* selected MEMs, thus representing the variation in community structured explained by combinations of selected geographical variables: (A) RDA1 for the FAPROTAX trait data; (B) – RDA1 for the distance-based FAPROTAX trait data. Positive values are represented by filled circles and negative values by open circles. Their sizes illustrate their absolute values, and their positions refer to the geographical position of each sampled shallow pond.

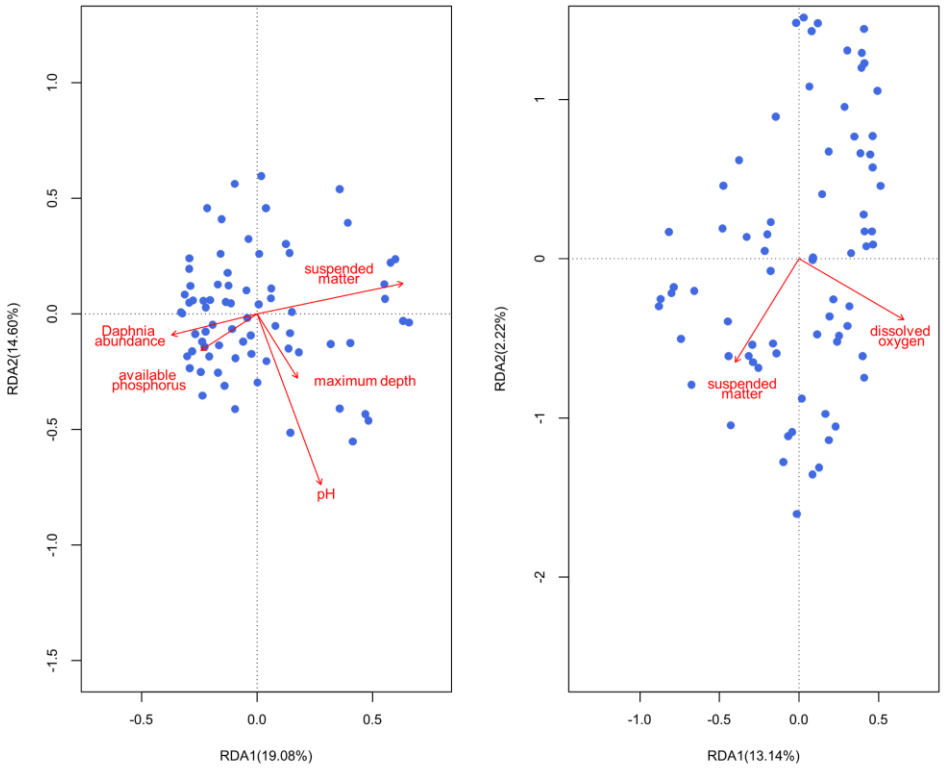

Figure S8. Redundancy analysis plot with environmental variables selected by forward selection for: (A) distance-based RDA on OTU data – each dot depicts a bacterioplankton community; (B) distance-based RDA on FAPROTAX data.

Chapter V

Bacterioplankton assembly is structured by zooplankton in a mesocosm experiment mimicking eutrophication and warming

Fabio Toshiro T. Hanashiro[1], Caroline Souffreau[1], Matthias Vanhamel[1], Edwin Van den Berg[1], Shinjini Mukherjee[2] & Luc De Meester[1,3,4]

[1]Laboratory of Aquatic Ecology, Evolution & Conservation, KU Leuven, Charles Deberiotstraat 32, B-3000 Leuven, Belgium

[2]Laboratory of Reproductive Genomics, KU Leuven, ON I Herestraat 49, B-3000 Leuven, Belgium

[3]Leibniz Institut für Gewässerökologie und Binnenfischerei (IGB), Müggelseedamm 310, 12587 Berlin, Germany

[4]Institute of Biology, Freie Universität Berlin, Königin-Luise-Strasse 1-3, 14195 Berlin, Germany

Publication status: in preparation

5.1 Abstract

Environmental selection, stochasticity and dispersal are known as important processes structuring bacterioplankton assembly, but the relative importance of these processes can be modified by climate warming and eutrophication. Here, we used outdoor mesocosms simulating shallow pond environments to quantify bacterioplankton assembly processes under a factorial combination of three different levels of eutrophication (nutrient enrichment), two temperature levels and a zooplankton richness gradient (16 different communities). Over a 12 weeks period, treatments led to strong environmental differences, ranging from oligotrophic to hypereutrophic conditions. However, these strong environmental differences did not translate to direct changes in bacterioplankton community composition and diversity. Different biotic factors were important to explain variation in bacterioplankton community composition, including zooplankton biomass and abundance of *Daphnia magna*. The most divergent communities at the final time point were mainly correlated with the presence of Cyanobacteria. Network analysis showed that the phylum Actinobacteria was negatively correlated with Cyanobacteria. This co-exclusion of specific phyla could be related to resource competition. Selective predation by zooplankton and zooplankton microbiome interactions might also be important in driving bacterioplankton variation. Our results emphasize the importance of biotic factors in structuring bacterioplankton communities. Furthermore, our findings highlight how different layers of complexity such as water chemistry (pH) and biotic interactions across multiple trophic levels are playing a role in bacterioplankton assembly under eutrophication scenarios.

Keywords: freshwater bacterioplankton, warming, eutrophication, biotic interactions, community assembly

5.2 Introduction

Bacterial communities play a pivotal role in different ecological functions across terrestrial and aquatic ecosystems, e.g., in nutrient recycling (Falkowski et al., 2008; Buchan et al., 2014; Martiny et al., 2015; Fierer, 2017). Therefore, determining which processes are driving their assembly is crucial, especially with the profound environmental changes on the planet going on. Global mean temperatures have already increased with 0.99°C since 1850-1900 (IPCC et al., 2021) and are expected to rise between 0.3 to 4.8 ºC in the next 100 years under IPCC scenarios (Stocker et al., 2013). Freshwater ecosystems are particularly impacted by climate change, directly by an increase in water temperature (O' Reilly et al., 2015) and indirectly by a combined effect with eutrophication leading to cyanobacterial blooms (Kosten et al., 2012; Huisman et al., 2018). Among freshwater systems, shallow ponds are interesting and practical systems to study how bacterial communities are responding to environmental changes. Shallow ponds are known to support high levels of biodiversity despite their size (De Meester et al., 2005; Beklioğlu et al., 2016), are particularly affected by pressures related to their surroundings (Scheffer & Van Nes, 2007; Brans et al., 2018), and are isolated "islands" within a terrestrial matrix which make them ideal systems to study assembly mechanisms that consider spatial signals as proxies for dispersal limitation in a metacommunity perspective (Souffreau et al., 2015, 2018; Hanashiro et al., 2019). Bacterioplankton communities in shallow ponds are also commonly used as ecological indicators due to their key roles in aquatic food webs and recycling of nutrients and organic matter (Logue & Lindström, 2008; De Bie et al., 2012; Souffreau et al., 2015; Muscarella et al., 2016). Understanding which processes underlie bacterioplankton assembly in shallow ponds could provide important subsidies on how to maintain their biodiversity and ecological functions under environmental change scenarios.

The mechanisms underlying bacterial community assembly are typically categorized as deterministic or stochastic (Zhou & Ning, 2017; Langenheder & Lindström, 2019). Deterministic processes encompass primarily niche-related mechanisms linked to local environmental selection (also known as species sorting in metacommunity ecology) by abiotic conditions (Van Der Gucht et al., 2007) or biotic interactions (Hanashiro et al., 2022). On the other hand, historical contingency (Rummens et al., 2018) and ecological drift (Zhou et al., 2013) are examples of stochastic processes potentially assembling bacterioplankton

communities. It is important to note that deterministic and stochastic processes are not mutually exclusive, and an interplay between both is probably the general rule (Langenheder & Székely, 2011; Stegen et al., 2013; Zhou & Ning, 2017; Langenheder & Lindström, 2019). The contribution of each type of process is dependent on other mechanisms such as legacy effects (Andersson et al., 2014), rarity (Liu et al., 2015) and habitat specialization (Székely & Langenheder, 2014) of the occurring species. Additionally, the relative importance of deterministic and stochastic processes can be modified by environmental changes (Langenheder & Lindström, 2019). Studies have shown that environmental change could lead to a reinforcement of environmental selection (e.g., acidification and desiccation) (Valverde et al., 2014; Tripathi et al., 2018). On the other hand, using an experimental approach, Zhang et al. (2016) showed that nutrient-enriched conditions increased the relative importance of stochastic over deterministic processes for soil communities. Similarly, bacterioplankton community assembly in mesocosms was mainly mediated by stochasticity when submitted to a combination of fertilization and warming (Ren et al., 2017).

Previous outdoor mesocosm studies explored the interaction effects of eutrophication and warming on bacterioplankton communities (Christoffersen et al., 2006; Shurin et al., 2012; Özen et al., 2013; Ren et al., 2017), however, little attention has been given to the effects of biotic interactions on bacterial structure, although these are fundamental to understand what drives bacterioplankton assembly, especially in scenarios of eutrophication. Zooplankton, for example, mediates ecological changes under eutrophic conditions directly as grazers on phytoplankton and bacterioplankton, and indirectly through alterations of environmental factors (Pernthaler, 2005; Scheffer & Van Nes, 2007; Huisman et al., 2018). Zooplankton species can also harbour high numbers of bacteria on their bodies and guts (Grossart et al., 2010; Eckert & Pernthaler, 2014). Qi et al. (2009) described close associations between three different zooplankton species of the genus *Daphnia* and bacterial taxa, suggesting a specific association between *Daphnia* and its bacterial symbionts. A certain degree of influence of zooplankton communities on bacterioplankton assembly could therefore be expected.

Here, we performed a mesocosm experiment and exposed bacterioplankton from shallow ponds to a combination of different temperature, nutrient-enrichment and zooplankton regimes to investigate their impact on bacterioplankton assembly. Our specific goals were (1) to evaluate the effects of warming and eutrophication on bacterioplankton assembly in a relatively natural setting of varying zooplankton richness and composition; and

(2) to evaluate to which degree abiotic and biotic factors influence bacterioplankton community composition under these anthropogenic gradients. We inoculated a mixture of natural bacterioplankton communities sampled from five shallow ponds located in Flanders, Belgium, to 80 mesocosms. Experimental treatments included two temperature levels (ambient and heated) and three levels of nutrient-enrichment using phosphorus and nitrogen (low nutrient level, high nutrient level, and increasing from low to high nutrient level). Temperature increase was only executed on the low nutrient and low-to-high nutrient treatments, resulting in total in five treatments. For each treatment, 16 different zooplankton communities were inoculated to the mesocosms, creating a gradient in zooplankton richness within each treatment. The key cladoceran species *Daphnia magna* was present in 8 out of 16 communities. Since we used different zooplankton communities, our outdoor mesocosms can be considered as individual shallow pond systems, each under different scenarios of nutrient-enrichment and temperature. In this way, our mesocosms simulated different scenarios along a eutrophication gradient including the biotic interactions mediated by the zooplankton community. After 12 weeks, the changes in bacterioplankton diversity and community composition (at different taxonomic levels) among the different environmental conditions were evaluated, and relations with abiotic and biotic conditions were assessed. We expected that bacterioplankton community structure would be strongly driven by the abiotic conditions created by the nutrient-enrichment and temperature treatments, but that the biotic interactions with zooplankton, and especially with the key grazer *Daphnia magna* would alleviate the effects of eutrophication. In contrast to a field survey, the outdoor mesocosms all started with similar founding bacterioplankton communities, and this way our study also provides insights about bacterioplankton community changes over time.

5.3 Material and methods

Experimental setup

Mesocosms were placed throughout summer (June – August 2014) at Arena area at KU Leuven University campus in a random position (Supplementary Fig. S1). On 1st of June 2014, these mesocosms were each filled with 170 L of tap water and 25 L of a standardized mixture of bacterioplankton and phytoplankton. This mixture was obtained by pooling water sampled from five different ponds located within 10 km of Leuven, Belgium (Supplementary Material

– Table S1). These ponds were selected due to the proximity to Leuven and because prior environmental data were available from previous projects (Gianuca et al., 2018; Verbeek et al., 2018). The pooled water samples were filtered through a 40 µm mesh plankton sieve to remove zooplankton and macroinvertebrates before adding them to the mesocosms. To obtain 16 different natural zooplankton communities over a gradient in species richness, we sampled 16 additional ponds (Table S1). Due to the known key role of *Daphnia magna* in structuring bacterioplankton communities by predation (Degans & De Meester, 2002) and symbiotic relationships (Sison-Mangus et al., 2015; Callens et al., 2016), half of these 16 ponds were chosen based on the presence of *Daphnia magna* (Table S1). Zooplankton communities were sampled using a 64 µm plankton net. Debris, aquatic plants and macroinvertebrates were manually removed, and an aliquot of 1 L was added to each mesocosm on 5th of June 2014. In total, a combination of 16 zooplankton communities x 5 different treatments (temperature and nutrient-enrichment – see description below) was used (N = 80 mesocosms). Since we had 5 different treatments and 16 zooplankton communities, our experiment did not have biological replicates, and each mesocosm is considered as a unique representation of an individual pond across a metacommunity (Fig. 1). Mesocosms were kept closed with a 1 mm mesh size lid, thus being exposed to rain conditions and colonization by external communities by air/water dispersal at some level. Manipulation during the experimental period was performed avoiding possible contamination among mesocosms (e.g., use of distilled water to rinse equipment between sampling/measurements of biotic and abiotic parameters).

A schematic overview of the experimental design is given in Figure 1. After a two-week acclimatization period for bacterioplankton and phytoplankton, and nine days for the zooplankton, the mesocosms were exposed to factorial combinations of three different nutrient levels (low – **LL**, low increasing to high nutrient level over time – **LH**, and high nutrient concentration level – **HH**) and two different temperatures (ambient – **LL, LH, HH**, and ambient +4°C – **LL+T** and **LH+T**). High nutrient concentration treatment (**HH**) was not combined with increase in temperature, resulting in five different treatments in total (LL, LH, HH, LL+T, LH+T). Abbreviations presented here will be used throughout the text.

Nutrient enrichments were performed using a combination of phosphorus (KH_2PO_4) and nitrogen ($NaNO_3$). Every three days nutrients were measured (methods described below) and when necessary, added to maintain the following desired nutrient levels. Low nutrient

level (**LL** and **LL+T**) was kept constant at 0.4 mg L^{-1} phosphorus and 3.2 mg L^{-1} nitrogen (1P:8N ratio) during the experiment. High nutrient level (**HH**) was kept at a concentration of 6.0 mg L^{-1} P and 48 mg L^{-1} N during the experiment. Low increasing to high nutrient level treatment (**LH** and **LH+T**) was gradually increased from the low nutrient level (cf. **LL**) at the start of the experiment to the high nutrient level (cf. **HH**) by the end of the experiment. The increase was performed every three days, maintaining the 1P:8N ratio (e.g., an increase of 0.23 mg L^{-1} P and 1.723 mg L^{-1} N). For the temperature treatment, we used a combination of heating elements (Eheim Thermocontrol® 300W), sensors (SM 811/2m VDH®) and thermostats (ALFA 72 VDH®) to keep the heated mesocosms always 4 ºC higher (**LL+T** and **LH+T**) than the ambient temperature (**LL**, **LH** and **HH**). The experiment lasted 12 weeks, from 14th of June 2014 to 30th of August 2014.

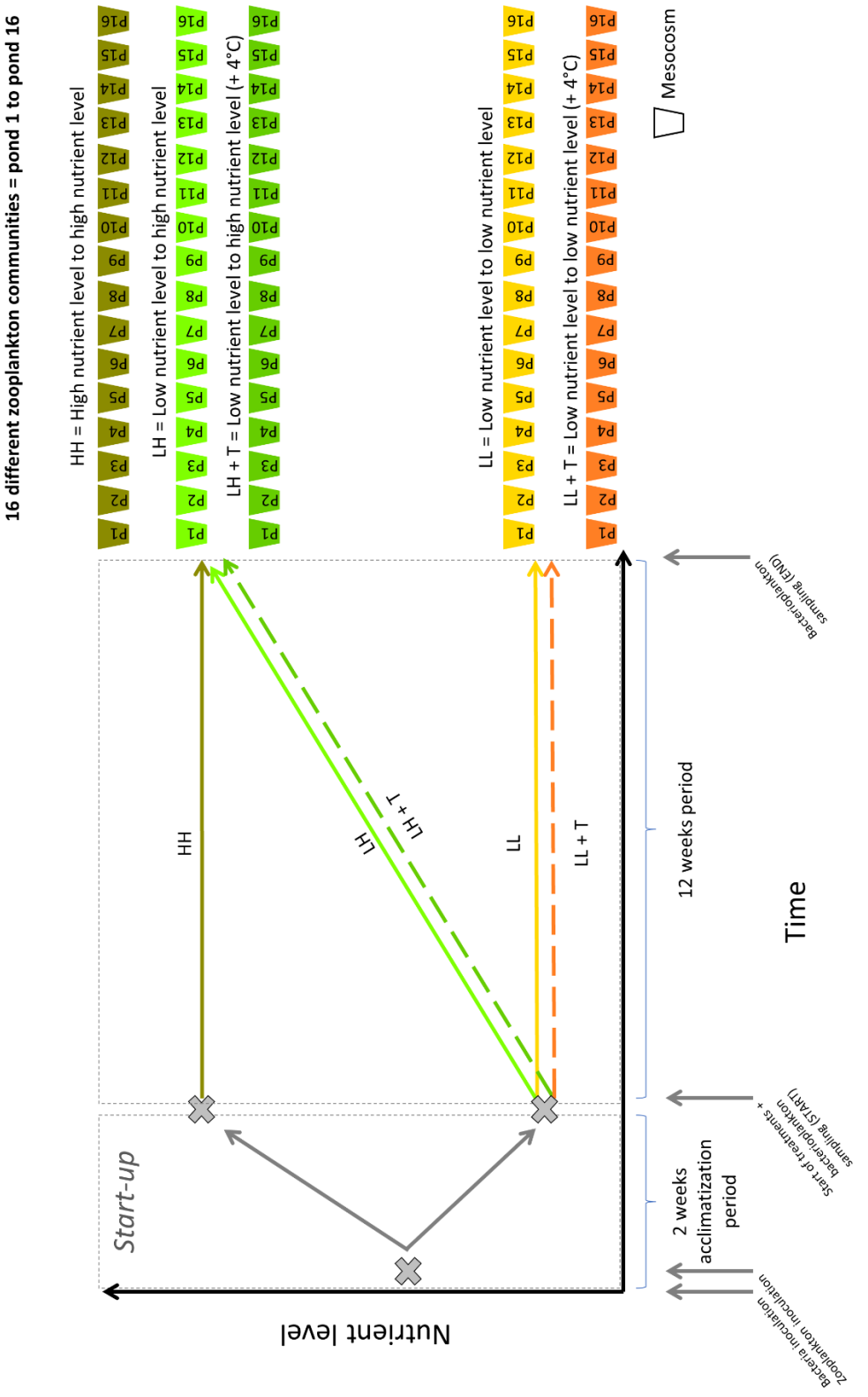

Figure 1. Schematic overview of the experimental design. Abbreviations for the nutrient treatments: (low – LL, low increasing to high – LH, and high nutrient concentration level – HH) and two different temperatures (ambient – LL, LH, HH, and ambient +4°C – LL+T and LH+T).

Measurement of biotic and abiotic parameters

Chlorophyll-a concentration, phycocyanin concentration and turbidity were quantified daily using a Turner Design AquaFluor® Handheld Fluorometer/Turbidimeter. At intervals of three days, pH, conductivity, dissolved oxygen and temperature were measured with a Hach® multimeter (United States). Every three days, samples to measure concentrations of total phosphorus and total nitrogen were taken by gently collecting 50 mL of surface water from each mesocosm. Total phosphorus concentration was estimated with the ascorbic acid method after perchlorate digestion (Murphy & Riley, 1962). Total nitrogen concentration was estimated using the Kjeldahl method (United States Environmental Protection Agency (USEPA), 1983). For reasons of feasibility, zooplankton was sampled two days before bacterioplankton communities at the end of the experiment, but also in intervals of three days throughout the 12 weeks experimental period. To sample zooplankton communities, 5 L of mixed water from the mesocosms were filtered through a 64 µm plankton net. Zooplankton samples were fixed with formalin (7% final concentration). For nine of the sampled time points, zooplankton was counted and identified with a stereomicroscope. Detailed information on zooplankton identification is given in Supplementary Material. From the zooplankton counts, the following explanatory variables were derived: abundances (individuals L^{-1}) for each zooplankton species, zooplankton biomass, *Daphnia magna* presence/absence, total zooplankton abundance, zooplankton richness, zooplankton Shannon's diversity, and the two first axes of a PCA (Principal Component Analysis) based on the log-transformed species abundance matrix. Diversity metrics and PCA axes were calculated only considering mature individuals (adults).

Bacterioplankton sampling, DNA extraction, amplification, sequencing and bioinformatics

Bacterioplankton communities were sampled on 14th of June 2014 at the start of the experiment (after two weeks of acclimatization), and on 30th of August 2014 at the end of the experiment. Water from the mesocosms was gently mixed before bacterioplankton sampling. We filtered up to 100 mL of water using a MCE 0.22 µm pore size filter (Milipore®, United States). Filters were stored at −80 ºC until DNA extraction. DNA was extracted from filters (no replicates per mesocosm) using the UltraCleanTM Soil DNA isolation kit (MoBio Laboratories, Inc., Carlsbad, CA, United States) following manufacturer's instructions. V3 and V4 regions (460 bp) of the 16S rRNA gene were amplified using the primers 16s-IllumTS-F

(CTGGAGTTCAGACGTGTGCTCTTCCGATCT – CCTACGGGNGGCWGCAG) and 16s-IllumTS-R (CTTTCCCTACACGACGCTCTTCCGATCT – GACTACHVGGGTAT CTAATCC) (Klindworth et al., 2013). Amplicon PCR was performed using a reaction mix containing 12.5 µL of Mytaq™ 2 x Master Mix, 2.5 µL of each primer (1 µM), 2.5 µL of DNA template (5 ng μL^{-1}) and 5 µl of molecular grade water. PCR temperature/time conditions were as follows: initial denaturation at 95 ºC for 60 s, 25 cycles of 15 s of denaturation at 95 ºC, 15 s of annealing at 55 ºC and 10 s of extension at 72 ºC. A final extension step was done at 72 ºC for 300 s. DNA purification was performed with Agencourt AMPure XP beads following manufacturer's instructions. An indexing PCR was subsequently conducted for sample identification using unique primer combinations of two different forward and 96 different reverse barcodes (total = 192 possible combinations). This nucleotide barcode dual-indexing system was performed following Lange et al. (2014). Indexing PCR was performed using a reaction mix of 10 µL of Mytaq™ 2 x Master Mix, 0.5 µL of forward primer (5 µM) and 1 µL of reverse primer (2.5 µM) and 8.5 µL of purified amplicon PCR product. Indexing PCR started with 60 s of initial denaturation at 95 ºC, followed by 12 cycles of 15 s of denaturation at 95 ºC, 15 s of annealing at 51 ºC and 10 s of extension at 72 ºC. A final extension step was performed for 300 s at 72 ºC. Products of indexing PCR were purified using Agencourt AMPure XP beads following manufacturer's instructions. PCR products were evaluated on gel electrophoresis and quantified using Quant-iT™ Picogreen® kit (Thermo Fisher). Samples with a minimum of 20 ng of amplicon DNA were pooled and diluted to final concentration of 10 nM and sent for sequencing on an Illumina MiSeq v3 PE300 (Genomics Core, KU Leuven, Belgium).

16S rRNA gene sequences were processed using the software QIIME (version 1.91) following the MiSeq standard operating procedure. 97% similarity level was used as operational taxonomic unit (OTU) definition. From the 160 samples, a total of 6,034,356 good-quality sequences were recovered. OTUs assigned to Archaea and OTUs with less than three reads were excluded from further analysis. Samples were rarefied to the same number of sequences (24,257sequences) to compare relative differences between samples.

Statistical analyses

All analyses were performed in R version 3.6.1 (R Core Team, 2018). We used richness (number of OTUs) and Shannon's diversity as indices for bacterioplankton taxonomic diversity. To compare taxonomic diversity across different mesocosms/treatments, we used Kruskal-Wallis

tests by ranks combined with Wilcoxon post-hoc test for pairwise comparisons, corrected by the Benjamini-Hochberg method for multiple testing. The same approach was used to compare environmental variable values across different mesocosms/treatments. For chlorophyll-a and phycocyanin concentration, and *Daphnia magna* abundances, values were log-transformed before the analysis. To visualize bacterioplankton community variation, ordination analysis (NMDS – Non-metric multidimensional scaling) was performed using Bray-Curtis dissimilarities as distance measure. To evaluate changes in bacterioplankton community composition across sampling times (start and end point), experimental treatments (5 treatments), and original zooplankton communities (with or without *Daphnia magna*; using pond identity as co-variable), we performed separate PERMANOVA analyses (999 permutations) using the function *adonis* of the package vegan (Oksanen, 2015). We used a CCA (Canonical Correspondence Analysis) combined with forward selection to assess the influence of local environmental variables driving changes on community composition.

To identify which OTUs were exclusively related to specific treatments and sampling times (start and end point), we used Venn diagrams that were generated with the function *venn* of the package venn (Dusa, 2021). To visualize the most common OTUs distribution across the different treatments, we used a heatmap generated with the function *pheatmap* of the package pheatmap (Kolde, 2019).

To better understand composition patterns of the bacterioplankton communities, we used a co-occurrence analysis based on pairwise correlations. The correlation matrix was calculated using all possible pairwise Spearman's rank correlations (> 0.6 or < -0.6) between the most abundant OTUs (relative abundance ≥ 0.1%). We used a p-value of < 0.05 and to reduce the rate of false-positive results, p-values were corrected using the Benjamini-Hochberg method. To visualize the correlation matrix generated, we used a network representation built in the interactive platform Gephi (Bastian et al., 2009). Correlation matrix calculation and network visualization were developed following analyses described in (Ju et al., 2014).

5.4 Results

Effect of nutrient-enrichment and warming on environmental conditions of the mesocosms

The five treatments led to differences in environmental conditions across the mesocosms at the final time point of the mesocosm experiment (Fig. 2). All significant pairwise comparisons are depicted in Fig. 2. Environmental variables known as proxies for eutrophication and productivity (chlorophyll-a concentration, phycocyanin fluorescence, turbidity, pH, conductivity and dissolved oxygen concentration) all had higher values for the nutrient-enrichment treatments. For chlorophyll-a concentration, HH and LH+T treatments had higher values compared to treatments without nutrient-enrichment (LL and LL+T) (global test – Kruskal-Wallis $\chi^2 = 22.42$, $p < 0.001$) (Fig. 2A) For phycocyanin fluorescence ($\chi^2 = 15.87$, $p = 0.003$) (Fig. 2B) and turbidity ($\chi^2 = 10.89$, $p = 0.027$) (Fig. 2C), only HH treatment had higher values compared to the LL+T treatment. pH was higher in nutrient-enrichment treatments (HH, LH and LH+T) compared to the treatments without nutrient enrichment (LL and LL+T) ($\chi^2 = 30.68$, $p < 0.001$) (Fig. 2D). Conductivity (Fig. 2E) and dissolved oxygen (Fig. 2F) followed the pattern of pH with significantly higher values in nutrient-enrichment treatments (HH, LH and LH+T) than in treatments without nutrient-enrichment (LL and LL+T) (Fig. 2D). Temperature, as expected, was higher in the two treatments with an increase in temperature (ambient + 4 ºC; LL+T and LH+T) ($\chi^2 = 28.81$, $p < 0.001$) (Fig. 2G). For the mesocosms in which *Daphnia magna* was present at the end of the experiment, *Daphnia magna* abundances were not significantly affected by the nutrient-enrichment and temperature treatments ($\chi^2 = 6.11$, $p = 0.1904$) (Fig. 2H). Time series of chlorophyll-a concentration, phycocyanin fluorescence, turbidity, pH, conductivity, dissolved oxygen concentration, temperature and *Daphnia magna* abundances are given in Supplementary Material (Fig. S2).

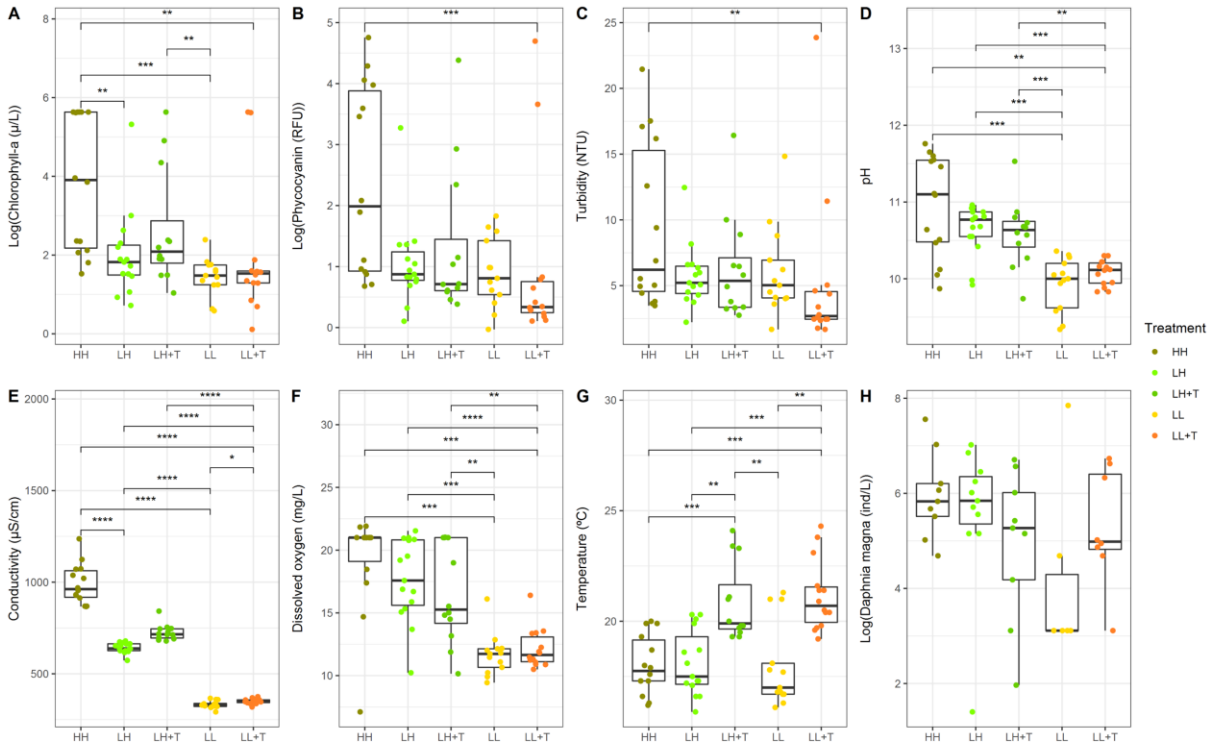

Figure 2. Boxplots of the measured environmental variables (A – chlorophyll-a concentration; B – phycocyanin fluorescence; C – turbidity; D – pH; E – conductivity; F – dissolved oxygen; G – temperature; H – *Daphnia magna* abundances) across the nutrient-enrichment and temperature treatments (see text for abbreviations) at the final time point. Significant pairwise comparisons are depicted in the figures with their respective significance levels (* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; **** $p < 0.0001$). RFU = Relative Fluorescence Units; NTU = Nephelometric turbidity unit.

Bacterioplankton community composition and diversity

Bacterioplankton community composition of the 80 mesocosm changed over time, with a significant difference between the start and end of the experiment (PERMANOVA – $R^2 = 0.326$, $p = 0.001$). Bacterioplankton communities showed a very similar composition at the start of the experiment (Fig. S3A). At the start, most OTUs were assigned to the phyla Proteobacteria (51.79%), Bacteroidetes (20.00%), Actinobacteria (8.38%) and Cyanobacteria (17.72%) (Fig. S3A). After 12 weeks, most OTUs were assigned to the phyla Actinobacteria (37.28%), Bacteroidetes (25.08%), Proteobacteria (24.43%) and Cyanobacteria (9.68%) (Fig. S3B). In order to test which phyla were responsible for the difference between start and final communities, a two-tailed t-test was performed for each phylum on its relative abundance data at start and end point, with Benjamini-Hochberg correction for multiple testing. Communities shifted from a dominance of Proteobacteria (*t-test* = -16.76, *adjp* = 0.001) at the

start of the experiment to a major prevalence of Actinobacteria (*t-test* = 10.83, *adjp* = 0.001) in the final time point of the experiment (Fig S3A and S3B).

Using hierarchical clustering and network analysis, we found an antagonism between the phylum Actinobacteria and Cyanobacteria in the final time point (Fig. 3A and Fig. 3B). Analysing the most abundant OTUs, hierarchical clustering shows that mesocosms with high abundances of Cyanobacteria had lower abundances of Actinobacteria (Fig. 2A). Network analysis shows that co-exclusions were mainly related to OTUs of Cyanobacteria (Chlamydomonaceae) and OTUs of Actinobacteria (Microbacteriaceae) (Fig. 3B, Fig. S4). The different OTUs of Cyanobacteria (Chlamydomonaceae) co-occurred altogether and with OTUs of the Proteobacteria (Erythrobacteraceae, *Polynucleobacter cosmopolitanus*, Xanthomonadaceae) (Fig. 3B, Fig. S4). One abundant OTU of the Bacteroidetes (*Leadbetterella*) formed a node with co-occurrences and co-exclusions with OTUs from the phylum Proteobacteria (Fig. 3B, Fig. S4).

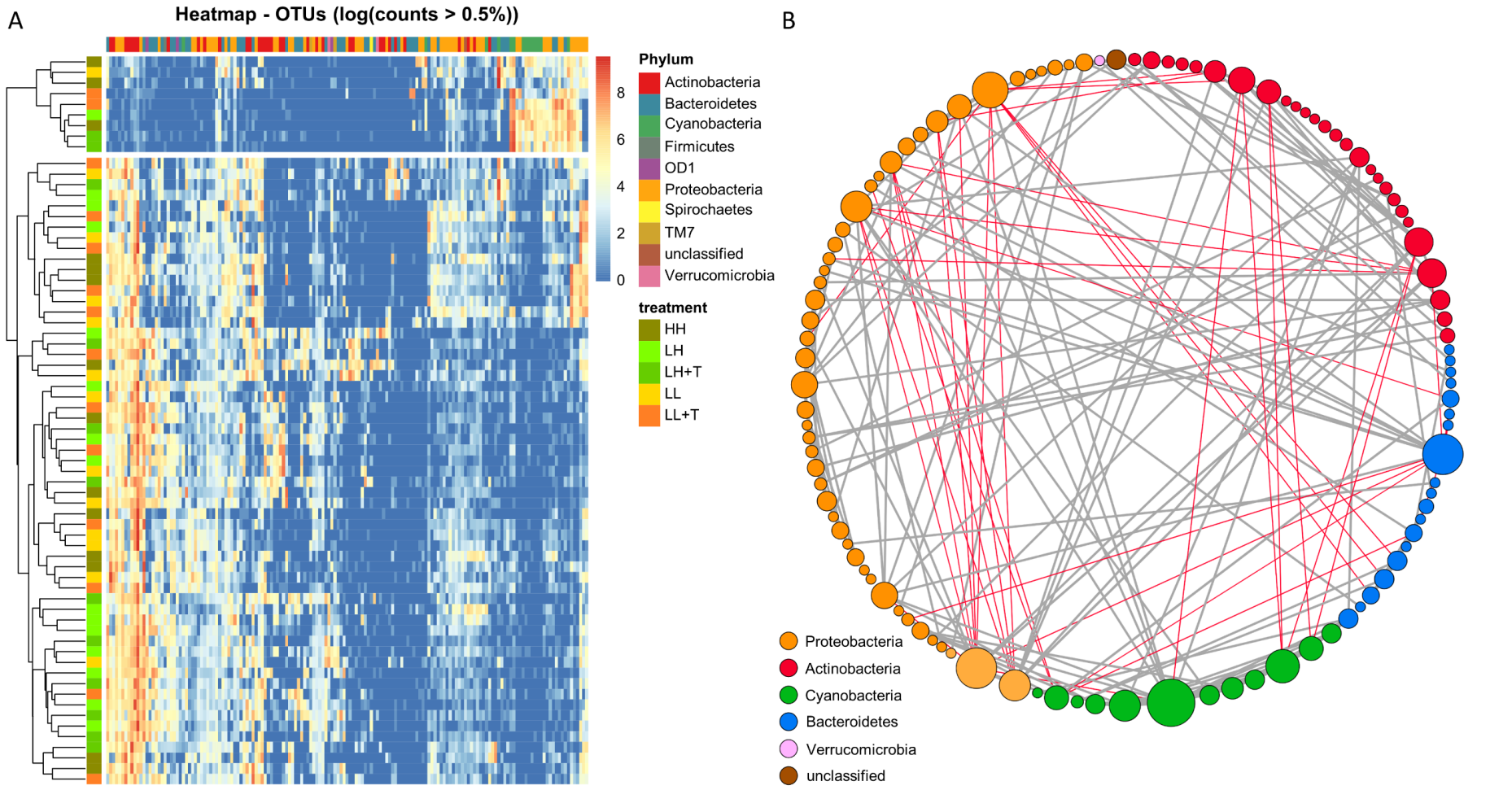

Figure 3. (A) Heatmap of the most abundant OTUs (relative abundance > 0.5%) from the different treatments at the final time point. OTUs are colored by their respective Phylum classification. Hierarchical clustering is based on the similarity of the different mesocosms across all treatments. (B) Network analysis of the OTUs with relative abundance higher than 0.1%. Edges stand for Spearman's correlation values (> 0.6, grey and < − 0.6, red) ($p < 0.05$). Node's sizes are proportional to the taxa mean relative abundance across samples. Nodes are colored by their respective Phylum classification.

No significant differences were found in terms of bacterioplankton diversity – estimated by the number of OTUs (Kruskal-Wallis, $\chi 2 = 1.08$, $p = 0.89$) and Shannon diversity ($\chi 2 = 1.30$, $p = 0.86$) – across the treatments at the final time point of the mesocosm experiment (Fig. S5). Among treatments of the final time point, 1354 OTUs (10.7% of the total number of OTUs) were shared among the five different treatments, with treatments HH and LH having higher number of exclusive OTUs, 1483 (11.8% of the total number of OTUs) and 1500 (11.9% of the total number of OTUs), respectively (Fig. 4B). Treatment LH+T had the lowest number of exclusive OTUs (990 – 7.8% of the total number of OTUs). Between time points, start samples had 16144 exclusive OTUs while final samples had 7598 exclusive OTUs (Fig. 4C).

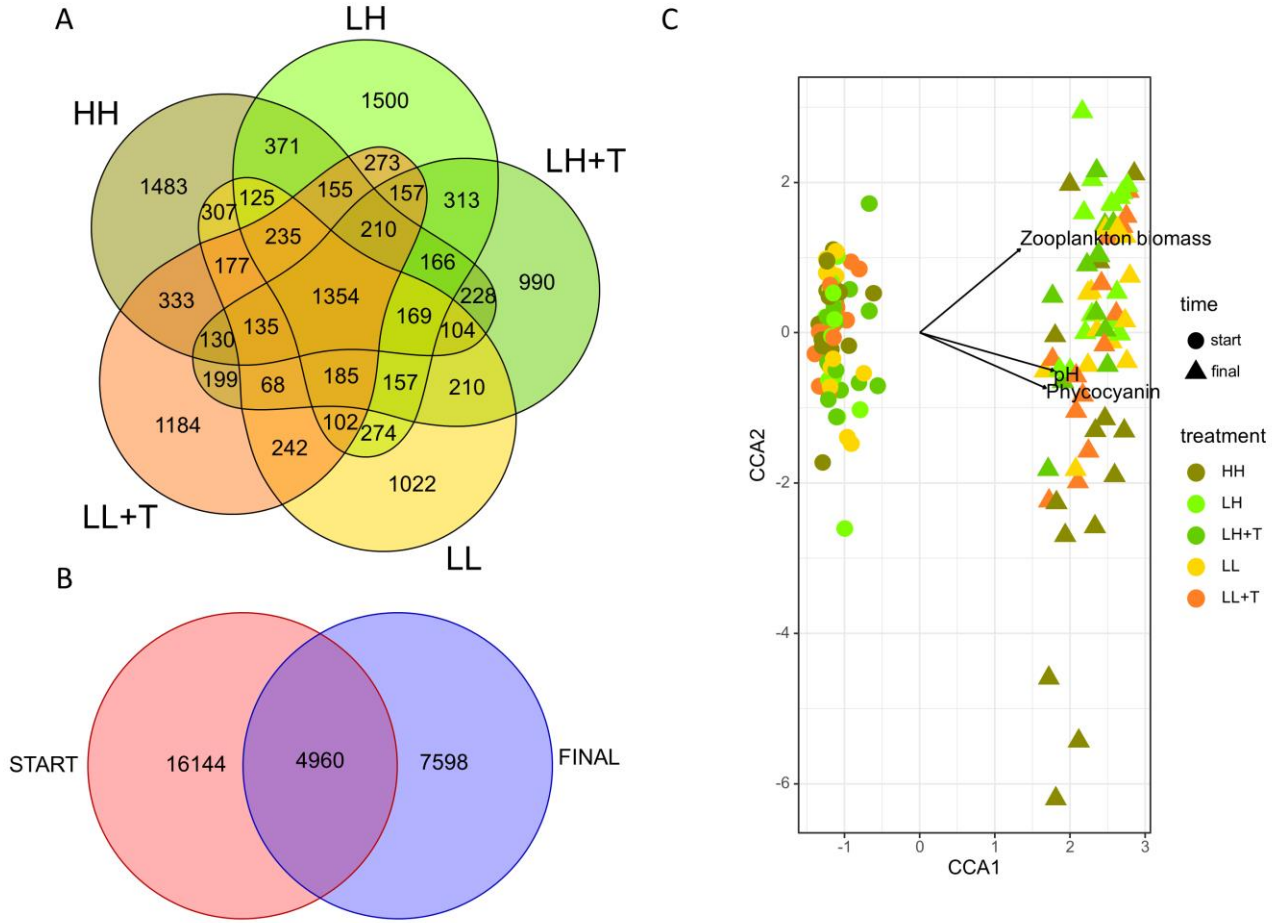

Figure 4. (A) Canonical Correspondence Analysis of bacterioplankton communities at the start and final time points of the experiment for all treatments, and the relationships with the environmental variables selected by forward selection. Communities are colored by treatment (see text for abbreviations). Venn diagrams showing the number of OTUs exclusive and shared by treatments in the final time point (B) and time points (C).

Environmental variables driving bacterioplankton community assembly

Despite the significant effects of treatment on the environmental variables (Fig. 2), no significant differences in bacterioplankton composition were found across the different treatments at the end of the experiment (PERMANOVA by treatment, –F = 0.63, p = 0.977) (see Table S2 for the main and interactive effects of the nutrient-enrichment and temperature factors). However, a significant effect of the original presence of *Daphnia magna* on bacterioplankton community composition was found (PERMANOVA, F = 6.93, p = 0.047) (Fig. S6) and an interaction effect between the pond identity of zooplankton communities (16 different communities) and the original presence of *Daphnia magna* on bacterioplankton community composition was found (PERMANOVA, F = 1.66, p = 0.048) (Fig. S6).

Pond identity of the zooplankton community did not significantly influence bacterioplankton richness (number of OTUs) (ANOVA, F = 1.54, p = 0.128) (Fig. S6A). However, a significant effect of the original presence of *Daphnia magna* on bacterioplankton diversity (number of OTUs) was observed (ANOVA, F = 5.56, p = 0.022) (Fig. S6B). For Shannon diversity, pond identity of the zooplankton community (ANOVA, F = 0.06, p = 0.799) (Fig. S7C) and original presence of *Daphnia magna* (ANOVA, F = 0.58, p = 0.863) (Fig. S7D) did not have a significant effect.

To get a better understanding of the possible effects of the measured environmental variables on community composition, we used Canonical Correspondence Analysis (CCA). The model generated by the CCA selected three variables driving changes in bacterioplankton community composition when incorporating both time points: pH (F = 5.44, p < 0.001), zooplankton biomass (F = 1.93, p < 0.001) and phycocyanin concentration (F = 1.16, p = 0.044) (Fig. 4A).

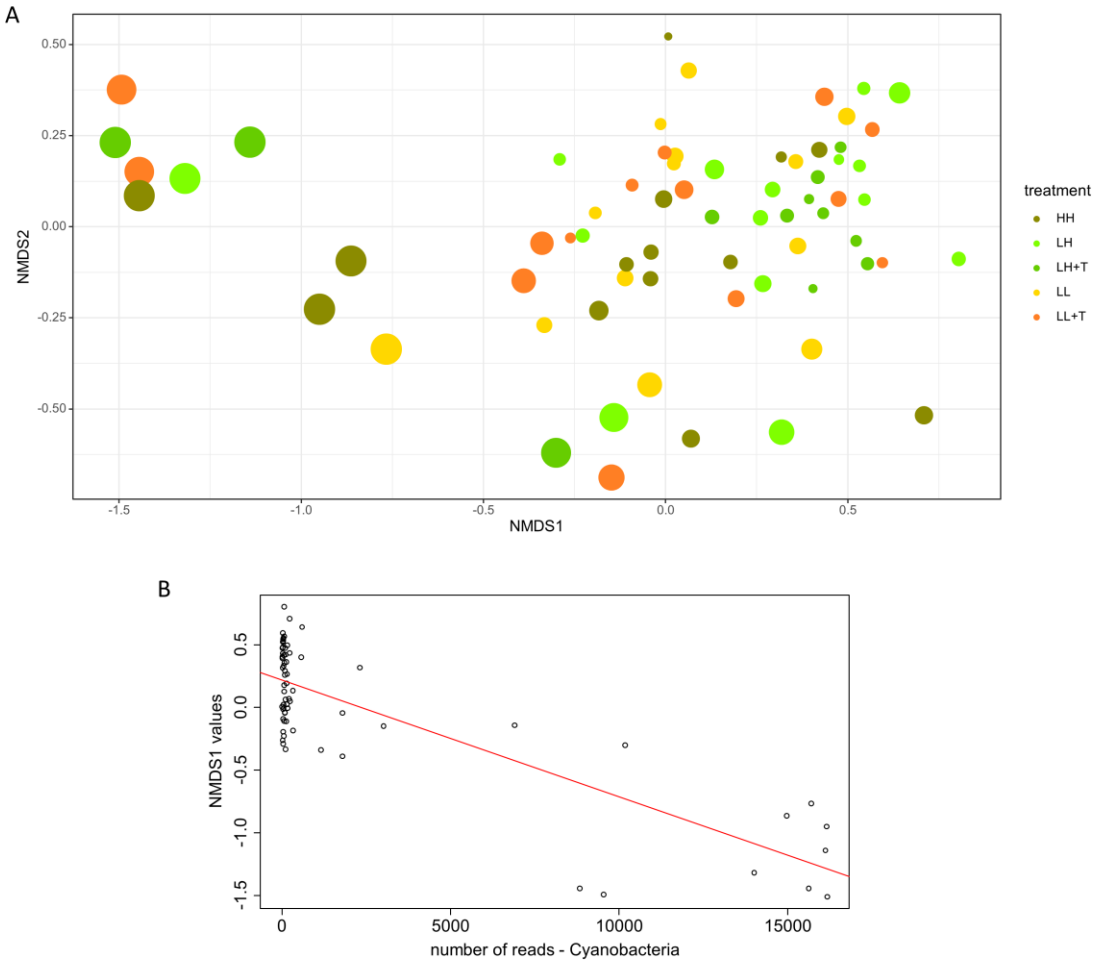

Figure 5. (A) Non-metric multidimensional scaling (NMDS) plot of bacterioplankton communities at the end point of the experiment. Communities are colored based on treatments. Circle sizes correspond to the number of reads of Cyanobacteria in the respective sample. (B) Correlation between the NMDS values (axis 1) and the number of reads of Cyanobacteria.

Analysing the final time point separately, no variables were selected by CCA. However, we found that communities with higher abundances of Cyanobacteria taxa were more dissimilar (Fig. 5A) (NMDS, STRESS = 0.15). A linear regression analysis between the NMDS values for the first axis and the number of reads of Cyanobacteria taxa present in the samples was significant ($adjR^2$ = 0.66, $p < 0.001$) (Fig. 5B).

Finally, in order to evaluate the role of environmental variables on specific bacterioplankton phyla, we performed Pearson's correlation tests with specific environmental variables at the final time point. We complemented the environmental variables selected in the previous analyses with abundances and biomass of individual zooplankton species, total zooplankton abundance and biomass, zooplankton diversity (richness and Shannon diversity)

and two first axes of a PCA (Principal Component Analysis) generated from the zooplankton abundance matrix. Regarding the most abundant phyla, Proteobacteria abundance was positively associated to *Daphnia pulex* biomass (r = 0.428, p < 0.01) and Bacteroidetes abundance was positively associated to cyclopoid copepod abundance (r = 0.37, p < 0.05). A complete description of the significant correlations is presented in Table S3. Only biotic variables were significantly correlated to relative abundances of the tested Phyla. Specific changes in bacterioplankton community composition among the treatments due to zooplankton abundance/biomass are represented for each of the 16 ponds as Supplementary Material (Figures S8 to S23).

5.5 Discussion

We used outdoor mesocosms under treatments of nutrient-enrichment, temperature increase and a zooplankton richness gradient to investigate changes in bacterioplankton assembly under varying environmental conditions. We were able to simulate different environmental conditions including a eutrophication scenario (from oligotrophic to hypereutrophic). Over the experimental period (12 weeks), all bacterioplankton communities diverged in composition compared to the initial inoculation (Fig. 4A). However, at the final time point, the nutrient-enrichment and temperature treatments did not lead to direct changes in community composition and diversity. Bacterioplankton assembly at the final time point was mainly governed by biotic factors related to the different zooplankton communities (total zooplankton biomass and presence of *Daphnia magna*). We also detected a pattern at the final time point where bacterioplankton community composition was partly structured by the relative abundance of Cyanobacteria (Fig. 5A), although we could not relate Cyanobacteria abundance to directly measured environmental variables such as nutrient concentration and temperature. Additionally, an increase in bacterioplankton richness related to the initial presence of *Daphnia magna* was reported. Interestingly, relative abundances of specific bacterioplankton phyla were correlated to the presence of specific zooplankton species.

Our experimental set-up was designed to represent the complexity of the natural world that might be underrepresented in mesocosms studies with more simple settings. Using 16 zooplankton communities combined with a zooplankton richness gradient, we aimed to create a more realistic metacommunity of small ponds under different conditions of nutrient

enrichment, temperature and biotic interactions. Bacterioplankton final community composition did not show a clear response to the strong changes in abiotic conditions (Fig. 2A-G), however, biotic interactions were related to most part of the changes in bacterioplankton community assembly. Along the two time points, the change in bacterioplankton community composition was partly related to an increase in total zooplankton biomass (Fig. 4A). The structuring role of zooplankton species on bacterioplankton assembly has been reported before with different processes being important, such as direct grazing (Berga et al., 2015) or changes in community composition (Verreydt et al., 2012; Birtel & Matthews, 2016; Hanashiro et al., 2019). In our study, probably the increase in total zooplankton biomass *per se* was not the responsible for driving changes on bacterioplankton assembly, but rather the growth of specific zooplankton species. At the final time point, *Daphnia magna* initial presence and abundance significantly affected bacterioplankton community structure (Figures S6 and S7). This Cladoceran is a zooplankton species well-known for its role in structuring bacterial communities through predation and symbiotic relationships with specific bacterial taxa (Souffreau et al., 2015; Callens et al., 2016). The structuring role of *Daphnia magna* on bacterioplankton structure could be explained by direct predation on bacteria itself, but also to indirect changes in environmental conditions e.g., due to phytoplankton grazing. Furthermore, other bacterioplankton phyla (Table S3) were significantly correlated to the densities of specific zooplankton taxa such as Cyclopoid copepods, *Scapholeberis mucronata*, *Simocephalus vetulus* and *Pleuroxus aduncus*. A positive relationship between the relative abundance of Proteobacteria and the biomass of *D. pulex* was found in our mesocosms (Table S3). Indeed, Proteobacteria taxa have been described as common *Daphnia* symbionts for different species (*D. magna*, *D. pulex* and *D. pulicaria*) (Qi et al., 2009). The increase in bacterioplankton richness in mesocosms where *D. magna* was initially present (Fig S6B) reinforce the idea of the structuring role of this cladoceran species on freshwater bacterioplankton communities. *Daphnia magna* could have controlled the dominance of phytoplankton species by grazing, allowing a higher number of bacterioplankton species to establish. These results highlight the necessity to include biotic factors (e.g., trophic interactions) in statistical models to better understand bacterioplankton assembly (Souffreau et al., 2015; Langenheder & Lindström, 2019; Hanashiro et al., 2022).

We also observed an interesting pattern related to Cyanobacteria in our study. At the final time point, bacterioplankton communities with higher numbers of Cyanobacteria taxa

seem to be more heterogeneous compared to communities with lower numbers of Cyanobacteria (Fig. 5A and 5B). Such divergent bacterial communities (higher beta-diversity) were also described in other mesocosms study and were there related to the effects of eutrophication (e.g., nutrient enrichment treatments), although in these cases not mediated through Cyanobacteria blooms (Ren et al., 2017). We visually detected this pattern for some of our mesocosms, with an increase in bacterial beta-diversity related to the nutrient-enrichment treatments (LH and HH), but not to the presence of high numbers of Cyanobacteria strains (Fig. 5A). Changes in diversity and composition of heterotrophic communities due to Cyanobacteria blooms have been described before with contrasting results. Due to the toxicity of some Cyanobacteria strains (e.g., *Microcysts*), blooms of this species are often associated with changes in bacterioplankton community structure by reducing its alpha diversity (Su et al., 2017). However, some organisms of the aquatic food web might be favored due to the production by Cyanobacteria of carbon and nitrogen compounds that can be consumed by higher trophic levels (e.g., protozoa, metazoan and heterotrophic bacteria) (Valverde et al., 2015). It is interesting to note that during Cyanobacteria blooms, environmental characteristics (e.g., higher dissolved organic content concentrations) were reported to be very similar among different ponds, but led to very dissimilar heterotrophic bacterial communities in these same ponds (Eiler & Bertilsson, 2004). It could be that effects on heterotrophic bacterial community variation might not be related to the environmental conditions created by the blooms, but to other factors such as biotic interactions between specific species through other ecological processes (e.g., competition for resources, toxins produced by Cyanobacteria).

Additionally, the relative abundance of Cyanobacteria was negatively related to that of Actinobacteria (Fig. 3A and 3B). This pattern has been reported before from metagenomes from around the world, indicating that Actinobacteria might be sensitive to the conditions leading to a cyanobacterial bloom (nutrient enrichment and high temperatures) (Ghai et al., 2014). Cyanobacteria strains could be more adapted to conditions of higher nutrient and organic matter availability, outcompeting Actinobacteria. We also observed a shift towards a higher contribution of Actinobacteria at the final time point compared to the initial dominance of Proteobacteria, possibly due to the increase of nutrients and warming (Fig. S3 and Fig. 2A). A similar response of bacterioplankton communities to nutrient-enrichment and warming was described in aquatic microcosms with a shift to a dominance of Actinobacteria (Ren et al.,

2013). The different community responses in some of our mesocosms might also be explained by a more tight relation between species adapted to eutrophication and its environmental conditions (Dai et al., 2017).

Even with biotic factors apparently being more important to determine changes on bacterioplankton community assembly in the final time point of our experiment, pH was also selected by our forward selection model as driving changes in community structure between the start and end point of the experiment. pH effects on bacterial community structure have been extensively described in different contexts (Griffiths et al., 2011; Ren et al., 2015), and are probably one of the most important structuring factors on bacterial assembly. Although we had significant differences in pH values among the treatments used in our study at the final time point, pH values were often alkaline since the start of the experiment (Fig. S2D) (> 9) and increased for all treatments in the final time point of the experiment.

While we expected to observe clear deterministic effects of nutrient-enrichment and warming on bacterioplankton in our mesocosm experiment, this was not the case. Previous experiments have reported contrasting results regarding the deterministic effects of nutrient-enrichment and warming on microbial community assembly. Ren et al. (2017) reported a combination of deterministic and stochastic processes structuring bacterial assembly under enhanced warming and nutrient-enriched conditions. In an another experiment, Zhou et al. (2014) studying the responses of groundwater microbial communities to nutrient-enrichment found that community succession was controlled mainly by stochastic processes. Our experimental setup aimed to create a metacommunity in which different factors were influencing bacterioplankton assembly. The addition of different zooplankton communities, combined with the zooplankton richness gradient added layers of complexity to the system, besides the (combination of) eutrophication and warming, and probably contributed to the apparent role of stochasticity in the bacterioplankton community assembly. Additionally, it could be that species adapted to eutrophic conditions were not present in the initial inoculum of our mesocosms, and that dispersal of such species from external sources (air, rain and soil) was only occasionally, resulting in a stochastic assembly for some mesocosms. Although stochastic processes are unpredictable by definition, factors and conditions leading to a stochastic assembly have been reported, such as: small habitat (Fukami, 2004), large species pool (Chase, 2003) and high productivity (Chase, 2010). Another possible explanation for the complex, relatively unpredictable community patterns found in our study is that the order and

success of species colonizing from outside the mesocosms might have influenced community dynamics more under hypertrophic conditions. This phenomenon is known as priority effects and has been described as a relevant mechanism for bacterial community assembly (Fukami et al., 2007, 2010). Also, although we measured multiple abiotic variables, we might have missed the exact variables driving bacterioplankton assembly in these specific conditions. Finally, the conditions created by the nutrient treatments were extreme in most of the mesocosms, going from eutrophic to hypereutrophic levels, and this might be an additional reason for the absence of clear patterns. Treatments as HH, where the nutrient level was kept in hypertrophic levels by adding inorganic phosphate and nitrate, probably exposed communities to conditions that led to high variability among mesocosms. These conditions probably drove changes in community structure and affected the trophic relations among bacteria, phytoplankton and zooplankton in the mesocosms. Alkaline conditions (high pH) as were registered in the mesocosms, combined with high photosynthesis activity was previously described changing bacterial activity in the carbon cycling (reducing biomass production and increasing C mineralization) (Tank et al., 2009). In our mesocosms carbon was not added and was therefore only available through phytoplankton biomass and excretion.

Community composition significantly changed from the start to the end of the mesocosm experiment, even for communities that were not treated with nutrient-enrichment or temperature increase (LL). This might be the attributed to seasonal variation during the summer period (Jones et al., 2012) or to a “mesocosm” effect (i.e. differences in ecological responses due to captivity), although microbial organisms tend to be less affected by the relatively small size of the mesocosms compared to larger organisms (Shade et al., 2018). Furthermore, communities under more “natural” conditions of nutrient levels and ambient temperature (LL – low nutrient level and ambient temperature) also showed differences in composition from the initial inoculum, which could be explained by seasonal variation or colonization by species from the air/soil or from other mesocosms (regional species pool) (Fig. 3A). To avoid creating gradients in space due to the daily sampling and manipulation of the mesocosms, they were randomly placed across the experimental area (Fig. S1). In order to check for a possible directional dispersal among spatially close mesocosms, we examined the NMDS plot using the original placement codes, but no clear pattern was found (Fig. S24). So, mesocosm manipulation and sampling did not lead to a directional dispersal, but we cannot

exclude an exchange by air/wind or by manipulation due to maintenance among mesocosms from the regional species pool.

Overall, we were able to demonstrate how under a combination of nutrient-enrichment, temperature increase and different zooplankton communities, bacterioplankton assembly could mainly be explained by biotic interactions. Total zooplankton biomass and the presence of specific zooplankton species such as *Daphnia magna* were the predominant factors influencing bacterioplankton community composition and richness across different treatments (nutrient-enrichment and warming) and different zooplankton communities. This might be explained by varying bacterioplankton species' success rates related to: (1) competitiveness in specific environmental conditions (nutrient enrichment and/or high temperature) in which Cyanobacteria blooms occurred; and (2) presence of specific zooplankton species (e.g., *Daphnia magna*) that could have different structuring roles, like predation, symbiotic relationships with bacterioplankton species or mediating changes in the local environmental conditions. While the presence of different zooplankton communities (combined with a zooplankton richness gradient) resulted in a more realistic bacterial metacommunity under different stressors (nutrient enrichment and temperature rise), our study highlights the importance of exploring these layers of complexity while trying to mimic aquatic systems with mesocosms. Additionally, we provide a better understanding of the effects of eutrophication, Cyanobacteria blooms and zooplankton interactions on bacterioplankton community dynamics, which becomes urgently needed under future scenarios of environmental change.

Acknowledgements

This study was supported by the ERA-Net BiodivERsA project TIPPINGPOND (Belspo) and by KU Leuven projects PF/2010/07 and C16/2017/002. FTTH was granted by the program Science Without Borders – CNPq/Brazil (245968/2012-1). MVH was supported by a PhD scholarship of the Institute for the promotion of Innovation by Science and Technology in Flanders (IWT). We thank Jelena Pantel, Andros T. Gianuca, Eyerusalem Goitom, Koen Rummens, Jessie Engelen, Kristien Brans, Xavier Karreman, Laurens Kilsdonk, Carla Denis, Melissa Schepens and Geert Neyens for their contributions to the field work, laboratory analyses and to the development of this project

5.6 Supplementary Material for Chapter V

Selection and identification of zooplankton communities

Selection of the 16 zooplankton communities used in the mesocosms was based on the presence/absence of *Daphnia magna* and in a gradient of number of different Cladocera species. Eight ponds that originally contained *D. magna* and varied between one to seven different Cladocera species were selected. Eight ponds without *D. magna* and with a gradient of one to eight different Cladocera species completed the dataset. Zooplankton was identified using an Olympus ZSX12 stereomicroscope following Flössner et al. (2000). Samples were counted to 300 individuals (continued up to 400 individuals when new species were found in the last 100 individuals). Individuals of the genus *Daphnia*, *Alona*, *Alonella, Bosmina, Pleuroxus*, *Scapholeberis*, *Simocephalus* and *Chydorus* were identified to species level, while *Ceriodaphnia* was kept to genus level.

Table S1. Overview of the 21 ponds used to inoculate the mesocosms with bacterioplankton and phytoplankton (five ponds), and with zooplankton (16 ponds). Inoculum type for which the ponds were used (bacterioplankton and phytoplankton, or zooplankton) is indicated, together with geographical coordinates. For all ponds, the zooplankton species present at the moment of sampling are given. Ponds with *Daphnia magna* are shown in bold.

	Phytoplankton	Bacterioplankton	Zooplankton	*Daphnia magna*	*Daphnia longispina*	*Daphnia obtusa*	*Daphnia pulex*	*Alona quadrangularis*	*Alona rectangula*	*Bosmina longirostris*	*Camptocerus rectirostris*	*Ceriodaphnia sp.*	*Chydorus sphaericus*	*Eurycercus lamellatus*	*Leydigia quadrangularis*	*Macrothrix hirsuticomis*	*Pleuroxus aduncus*	*Pleuroxus uncinatus*	*Polyphemus pediculus*	*Scapholeberis mucronata*	*Scapholeberis ramneri*	*Simocephalus vetulus*	Cyclopoid copepods	Calanoid copepods	Zooplankton Richness
Pond ID & geographical coordinates	Inoculation type		Zooplankton species present																						
PL25-RED (50°47'3.1", 3°54'57.4")			x	x			x					x	x										x		4
AnOudIn (51°20'15.21", 5°2'21.74")			x	x	x		x					x	x										x		6
VBHalNa (50°47'17.43", 4°15'23.19")			x	x		x	x					x	x				x			x		x	x		9
PL1-GRE (50°51'49.04", 4°43'24.35")			x	x																			x		2
PL15-YEL (51°2'33.6", 3°43'50.02")			x	x	x		x						x				x				x		x		7
Mech-KT2 (51°01'26.1", 4°29'01.8")			x	x	x		x					x	x				x						x		7
VBDieIn (50°57'36.31", 4°58'20.50")			x	x	x	x	x						x							x		x	x	x	9
LRV (50°49'50", 4°38'15")	x	x	x	x	x		x		x			x	x							x		x	x	x	10
Mech-KT2 (51°01'26.1", 4°29'01.8")			x		x					x		x					x					x	x		7
PL3-YEL (50°50'35.43", 4°51'38.74")			x		x		x						x		x		x						x		6
PL18-GRE (51°10'41.2", 4°9'39.6")			x		x		x					x										x	x		5
PL18-RED (51°9'44.1", 4°9'15.1")			x		x								x										x		3
PL2-YEL (50°57'36.53", 4°44'22.54")			x				x			x		x	x				x					x	x		7
PL11-GRE (50°50'38.3", 3°36'57.1")			x		x							x	x									x	x		5
PL14-GRE (50°56'56.22", 3°43'12.37")			x																			x	x		2
PL14-YEL (50°57'36.16", 3°43'8.05")			x		x							x	x								x		x		5
Kessel-Lo (50°53'40", 4°43'24")	x	x			x			x						x									x	x	5
Abdijpark (50°51'48", 4°43'26")	x	x																							
Kasteelpark (50°51'42", 4°40'55")	x	x								x	x											x	x		4
Zoetwater (50°49'25", 4°39'50")	x	x			x			x		x		x	x	x		x	x	x	x	x		x	x		13

Table S2. PERMANOVA table of the effect of the nutrient-enrichment and temperature treatments on the bacterioplankton community composition at the final time point.

Source	df	SS	*F*	*p*
Nutrient	1	0.2185	0.6965	0.792
Temperature	1	0.2086	0.6653	0.848
Nutrient x Temperature	1	0.1719	0.5480	0.956
	64	20.0755		

Table S3. Pearson's correlations among taxa at Phylum level (relative abundance) and environmental variables. Significance values: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Taxa	Environmental variable	Correlation and significance
Bacteroidetes	Cyclopoid copepods	0.3747*
Chlamydiae	Axis (PCA2)	-0.3996**
OD1	*Scapholeberis mucronata*	0.4078**
Proteobacteria	Biomass of *Daphnia pulex*	0.4289**
Spirochaetes	Axis (PCA2)	-0.4513**
SR1	Biomass of *Simocephalus vetulus*	0.3733*
TM7	*Simocephalus vetulus*	0.3918*
Verrucomicrobia	Juvenile copepods	0.4476**
	Pleuroxus aduncus	0.5569***
	Scapholeberis_mucronata	0.5148**



Figure S1. Mesocosms at Arena area at KU Leuven University Arenberg campus.

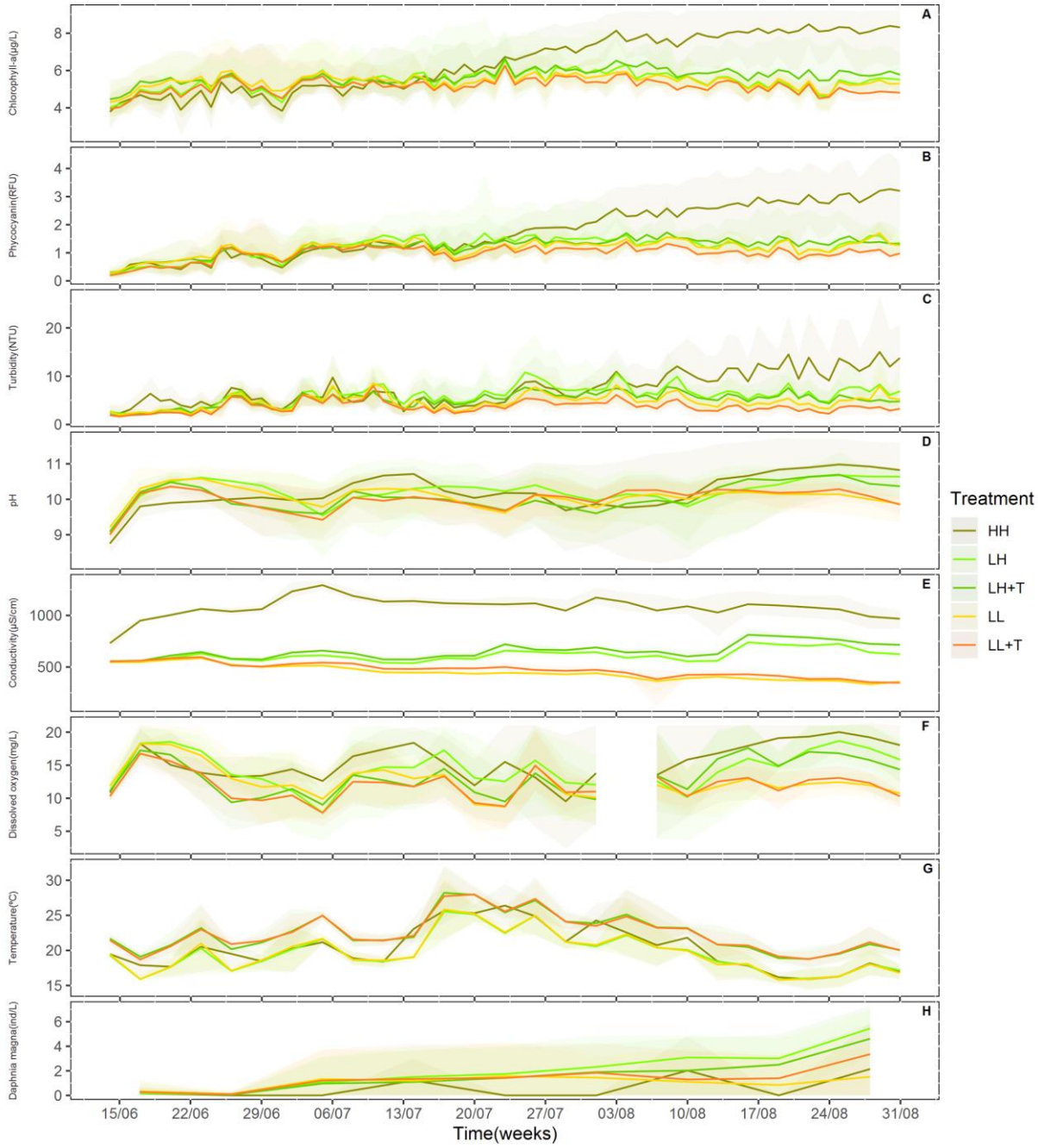

Figure S2. Daily values for chlorophyll-a concentration (A), phycocyanin fluorescence (B) and turbidity (C). Average three days interval values for pH (D), conductivity (E), dissolved oxygen concentration (F) and temperature (G). Nine time points for *Daphnia magna* abundances (H). Coloured lines depict the average values across mesocosms for the five treatments (HH, LH, LH+T, LL and LL+T) and shadow areas show lower (10%) and upper (90%) quantiles. Values for chlorophyll-a concentration (A), turbidity (C) and *Daphnia magna* abundances (H) are shown in logarithmic scale.

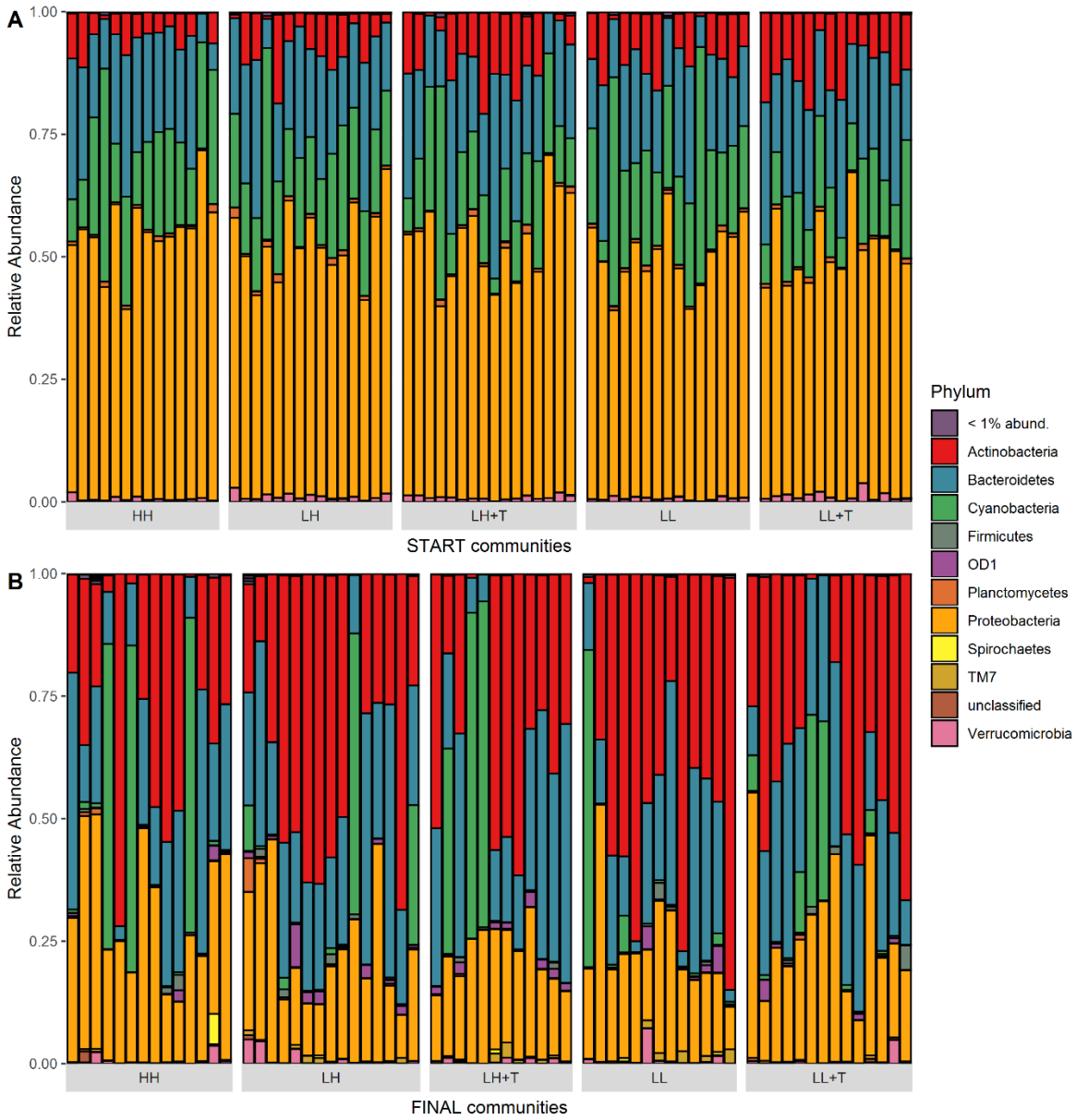

Figure S3. Relative abundance profiles of the dominant phyla present in the bacterioplankton communities at the start point (A) and final point (B) of the experiment. Communities are ordered by the five treatments (HH, LH, LH+T, LL and LL+T).

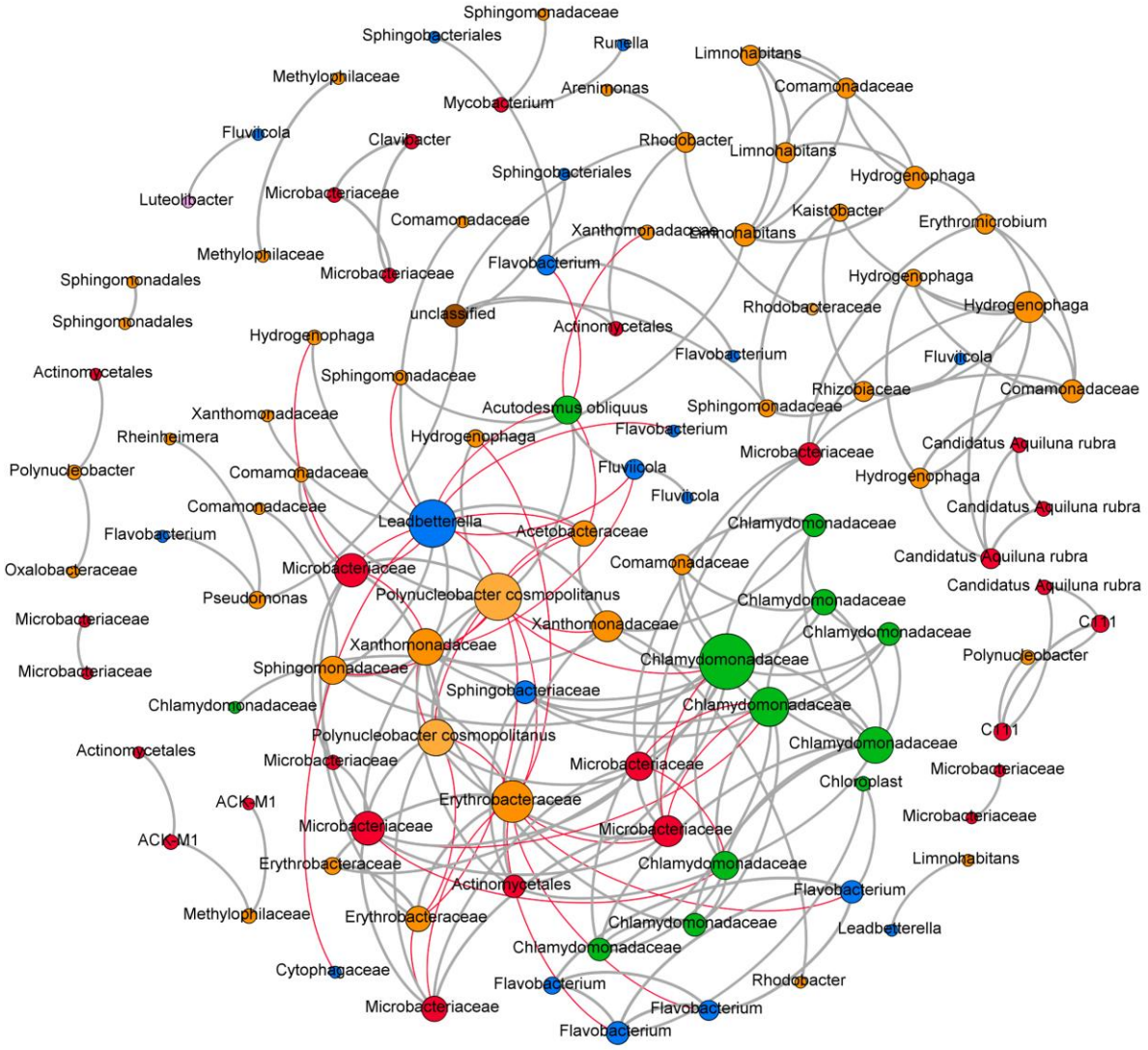

Figure S4. Network analysis of the OTUs with relative abundance higher than 0.1%. Edges stand for Spearman's correlation values (> 0.6, grey and < − 0.6, red) ($p < 0.05$). Node's sizes are proportional to the taxa mean relative abundance across samples. Nodes are colored by their respective Phylum classification and labelled with the lowest assigned taxonomic level.

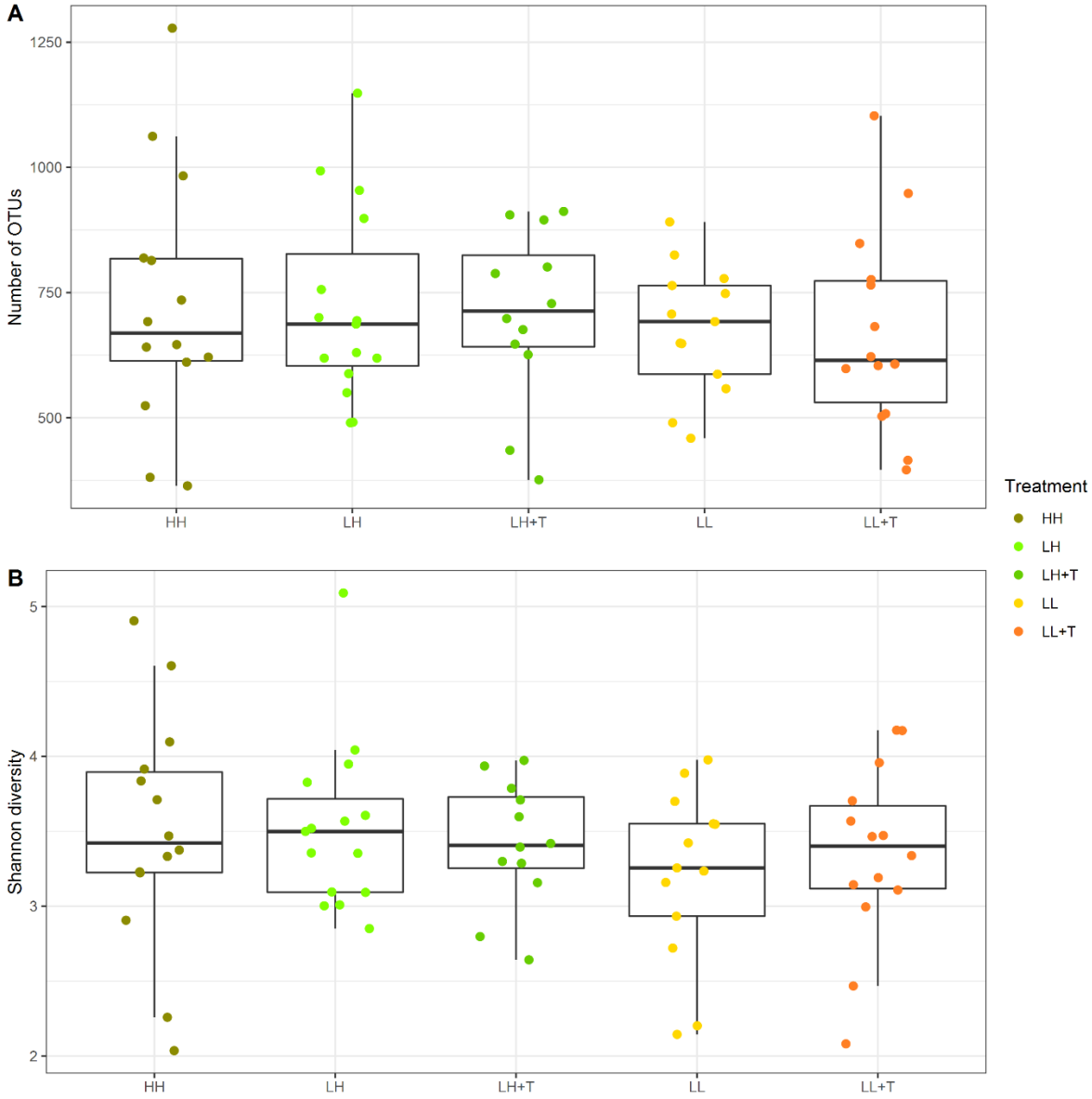

Figure S5. Bacterioplankton diversity boxplots – (A) number of OTUs and (B) Shannon diversity – across the five treatments (HH, LH, LH+T, LL and LL+T).

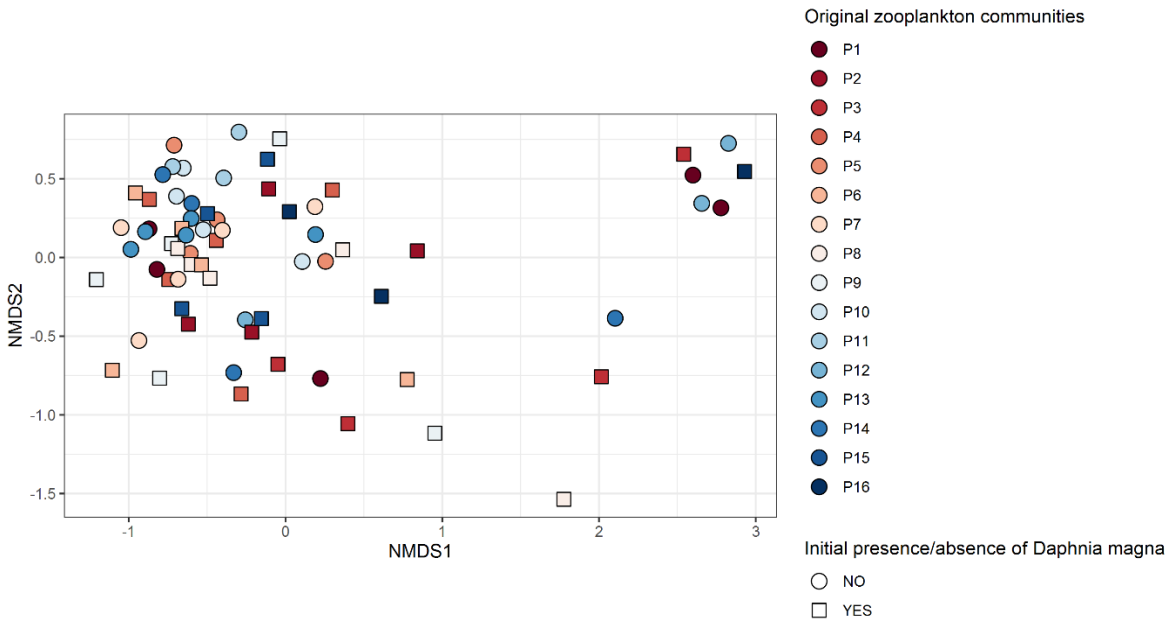

Figure S6. Non-metric multidimensional scaling (NMDS) plot of bacterioplankton communities at the final time point of the experiment. Colors and shapes are displaying the original zooplankton communities (16 levels) with and without *Daphnia magna*.

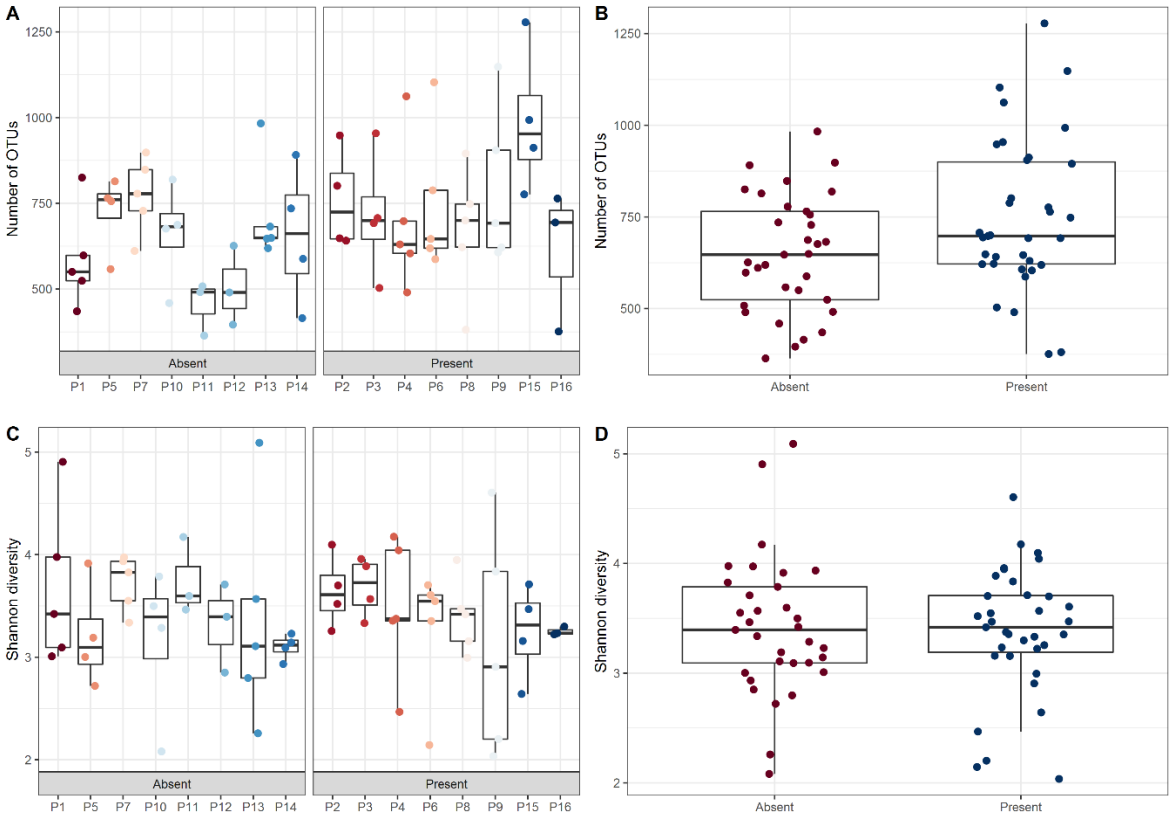

Figure S7. Bacterioplankton diversity boxplots – (A) number of OTUs across the 16 ponds with original absence/presence of *Daphnia magna*; (B) number of OTUs for original absence/presence of *Daphnia magna*; (C) Shannon diversity across the 16 ponds with original absence/presence of *Daphnia magna*; (D) Shannon diversity for original absence/presence of *Daphnia magna*

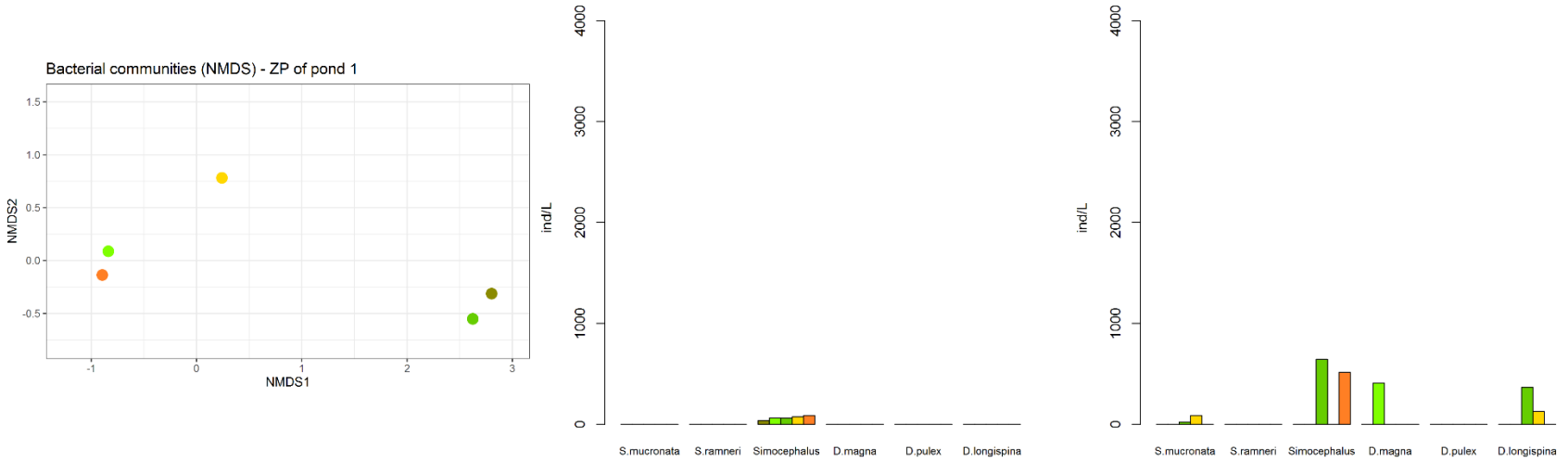

Figure S8. Non-metric multidimensional scaling (NMDS) plot of bacterioplankton communities at the end point of the experiment. These mesocosms were inoculated with zooplankton communities from pond 1 (see Table S1 for more details). Changes in abundance of the most important zooplankton species are shown in the barplots for both time points (start and final).

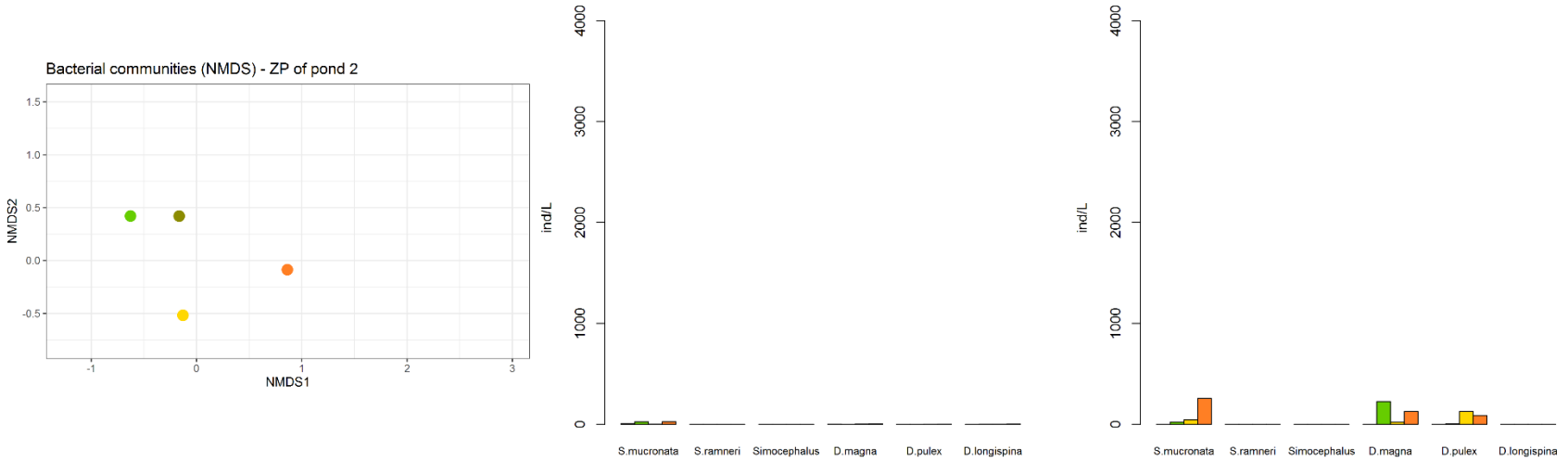

Figure S9. Non-metric multidimensional scaling (NMDS) plot of bacterioplankton communities at the end point of the experiment. These mesocosms were inoculated with zooplankton communities from pond 2 (see Table S1 for more details). Changes in abundance of the most important zooplankton species are shown in the barplots for both time points (start and final).

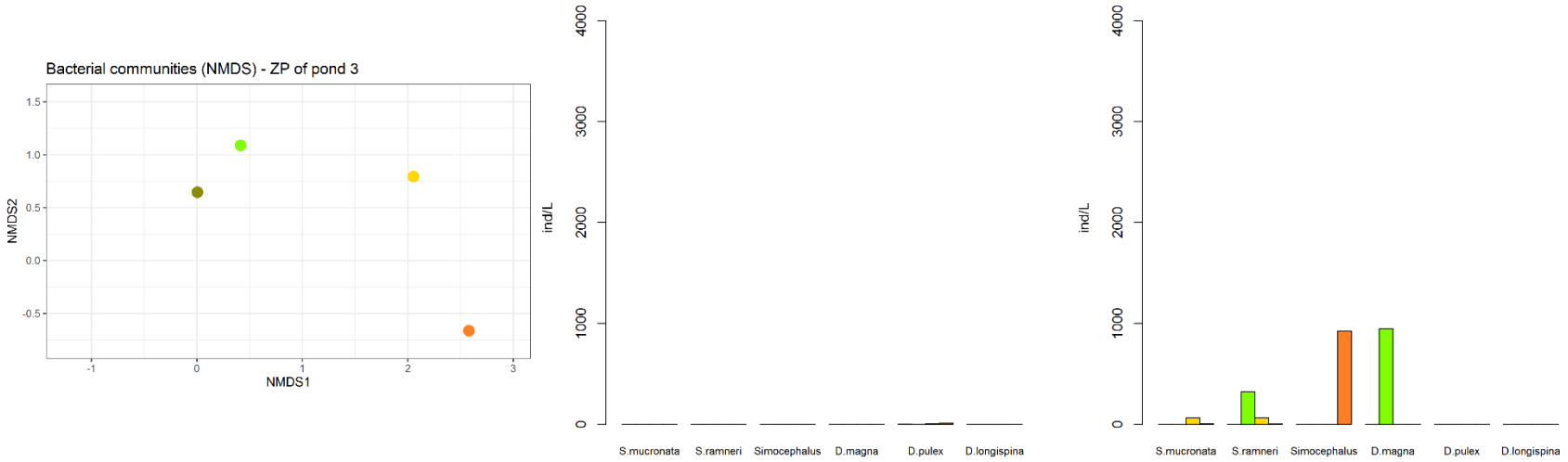

Figure S10. Non-metric multidimensional scaling (NMDS) plot of bacterioplankton communities at the end point of the experiment. These mesocosms were inoculated with zooplankton communities from pond 3 (see Table S1 for more details). Changes in abundance of the most important zooplankton species are shown in the barplots for both time points (start and final).

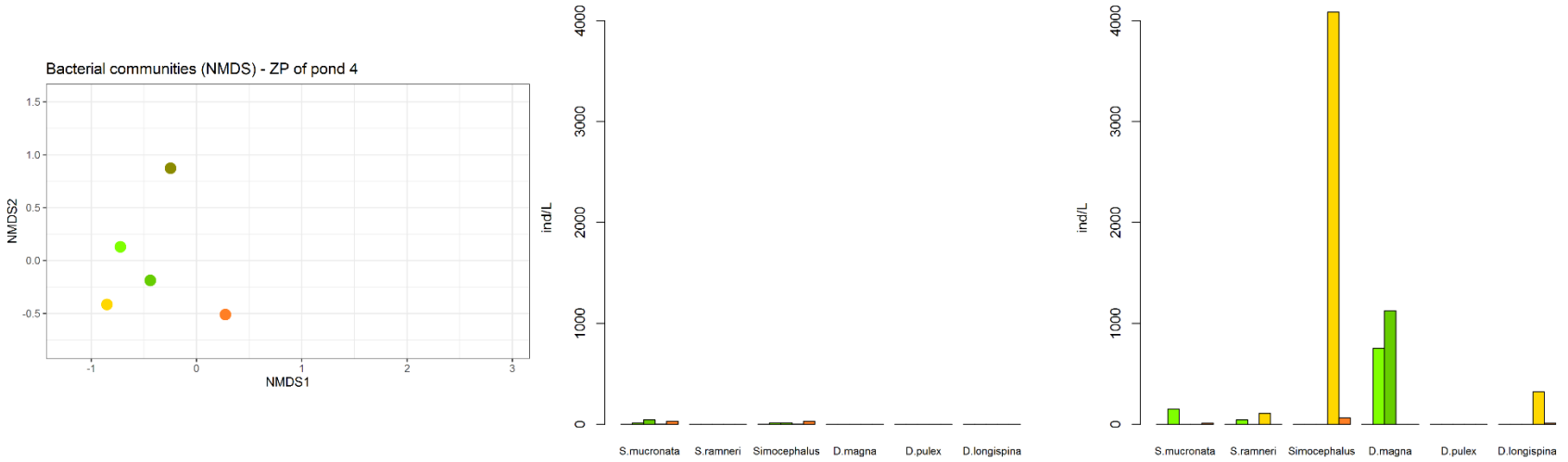

Figure S11. Non-metric multidimensional scaling (NMDS) plot of bacterioplankton communities at the end point of the experiment. These mesocosms were inoculated with zooplankton communities from pond 4 (see Table S1 for more details). Changes in abundance of the most important zooplankton species are shown in the barplots for both time points (start and final).

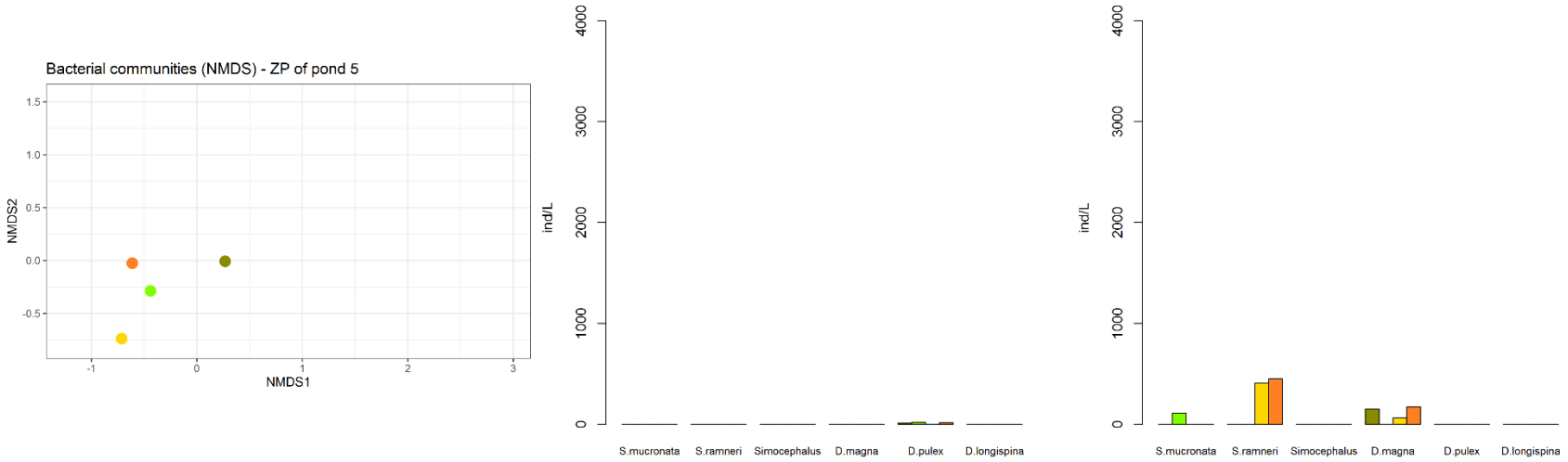

Figure S12. Non-metric multidimensional scaling (NMDS) plot of bacterioplankton communities at the end point of the experiment. These mesocosms were inoculated with zooplankton communities from pond 5 (see Table S1 for more details). Changes in abundance of the most important zooplankton species are shown in the barplots for both time points (start and final).

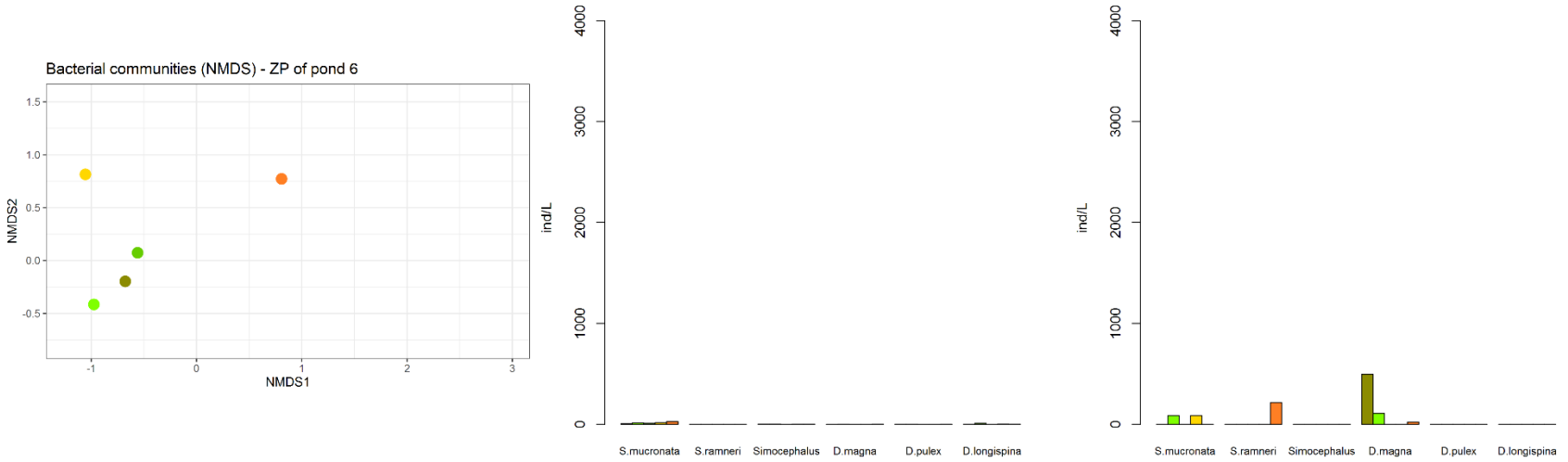

Figure S13. Non-metric multidimensional scaling (NMDS) plot of bacterioplankton communities at the end point of the experiment. These mesocosms were inoculated with zooplankton communities from pond 6 (see Table S1 for more details). Changes in abundance of the most important zooplankton species are shown in the barplots for both time points (start and final).

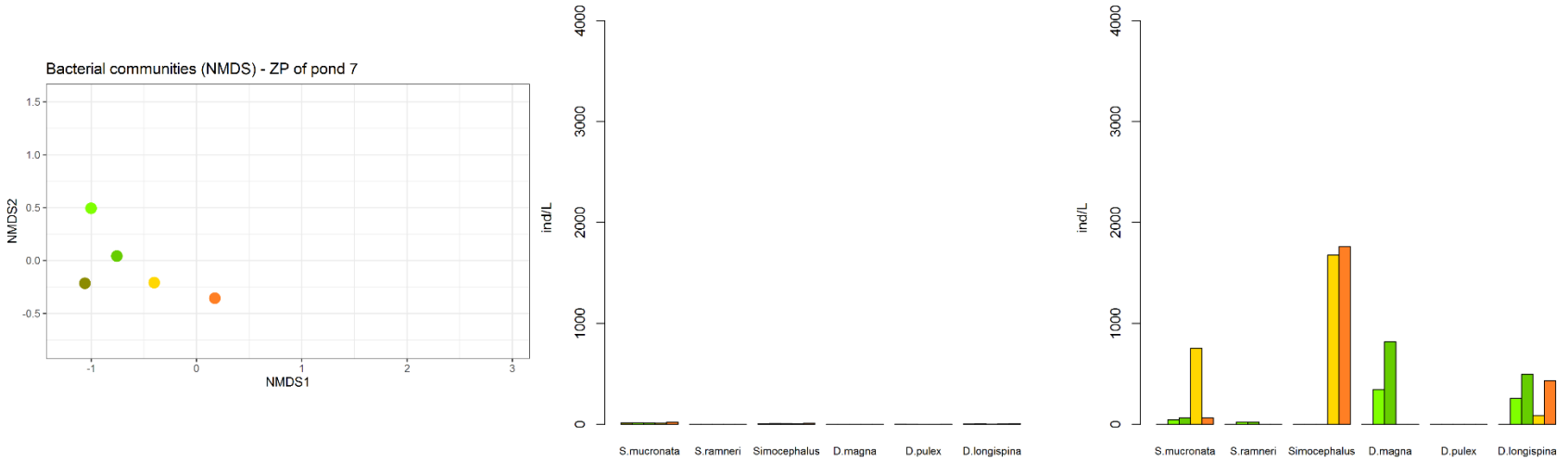

Figure S14. Non-metric multidimensional scaling (NMDS) plot of bacterioplankton communities at the end point of the experiment. These mesocosms were inoculated with zooplankton communities from pond 7 (see Table S1 for more details). Changes in abundance of the most important zooplankton species are shown in the barplots for both time points (start and final).

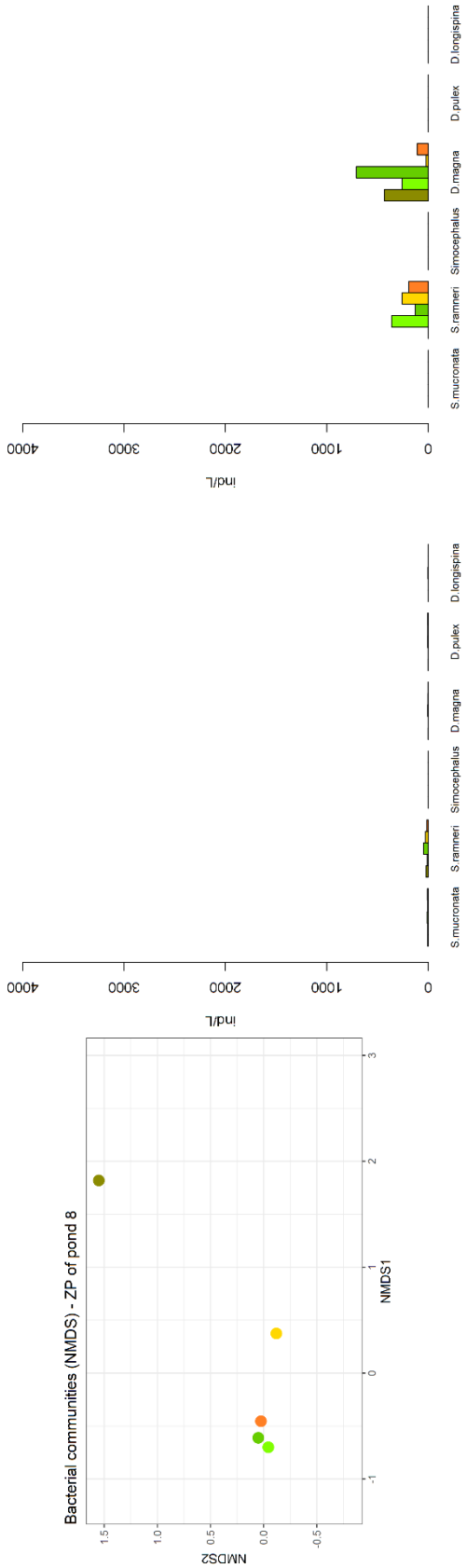

Figure S15. Non-metric multidimensional scaling (NMDS) plot of bacterioplankton communities at the end point of the experiment. These mesocosms were inoculated with zooplankton communities from pond 8 (see Table S1 for more details). Changes in abundance of the most important zooplankton species are shown in the barplots for both time points (start and final).

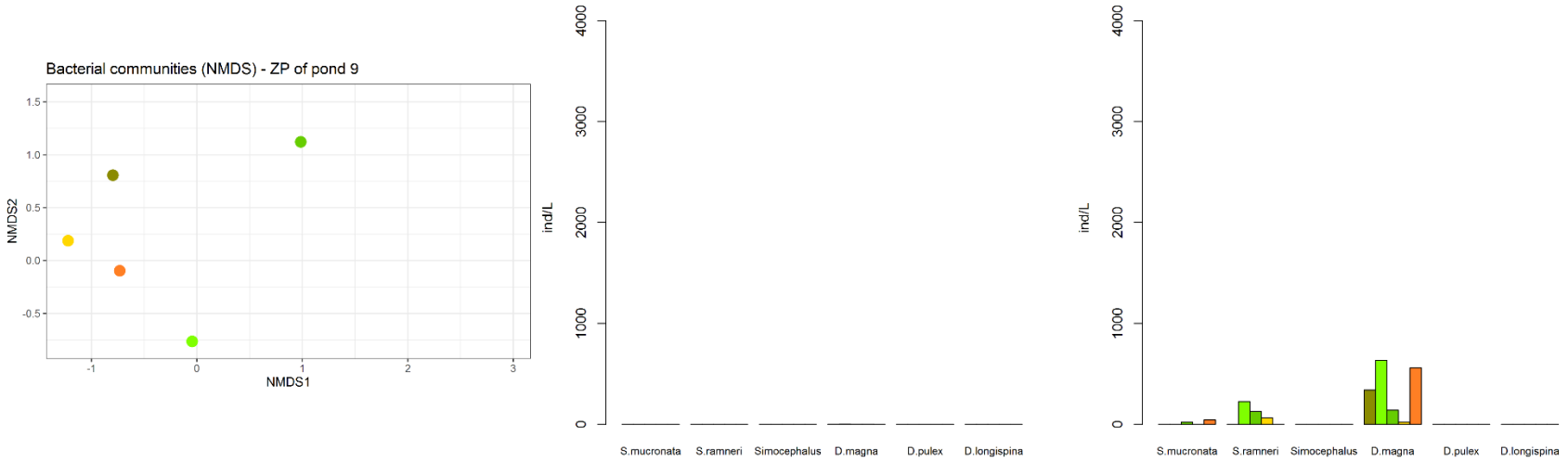

Figure S16. Non-metric multidimensional scaling (NMDS) plot of bacterioplankton communities at the end point of the experiment. These mesocosms were inoculated with zooplankton communities from pond 9 (see Table S1 for more details). Changes in abundance of the most important zooplankton species are shown in the barplots for both time points (start and final).

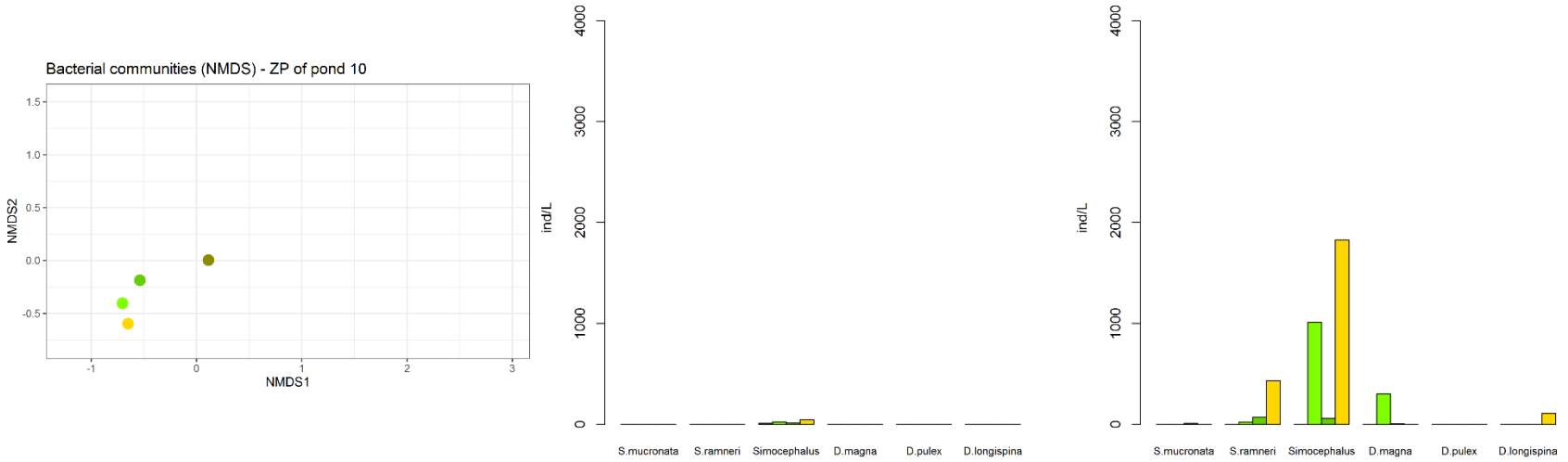

Figure S17. Non-metric multidimensional scaling (NMDS) plot of bacterioplankton communities at the end point of the experiment. These mesocosms were inoculated with zooplankton communities from pond 10 (see Table S1 for more details). Changes in abundance of the most important zooplankton species are shown in the barplots for both time points (start and final).

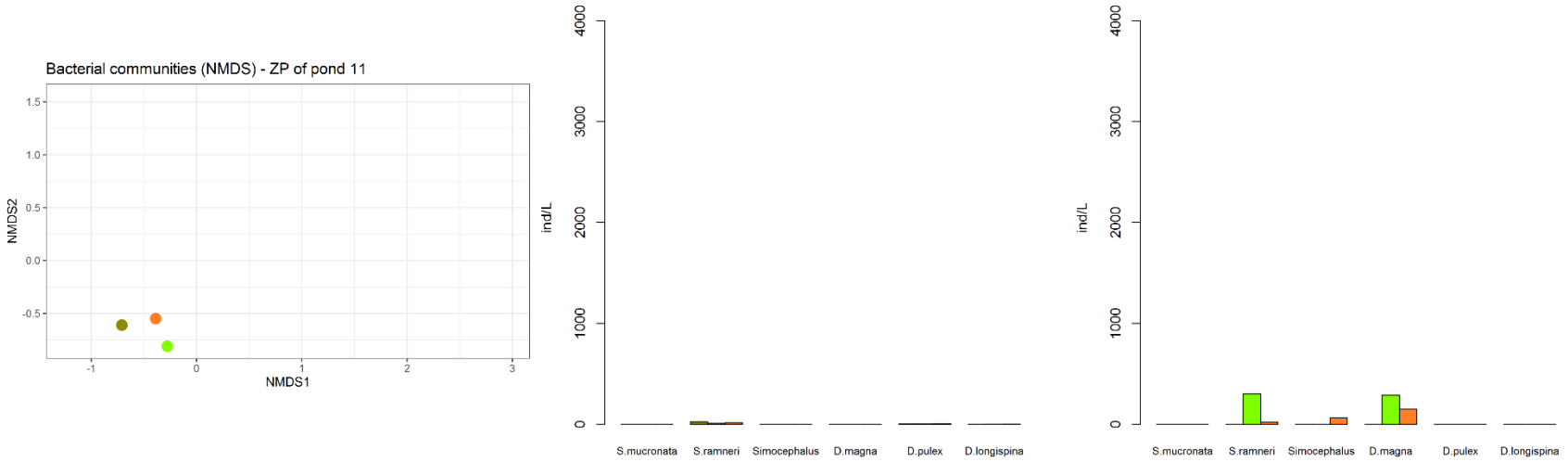

Figure S18. Non-metric multidimensional scaling (NMDS) plot of bacterioplankton communities at the end point of the experiment. These mesocosms were inoculated with zooplankton communities from pond 11 (see Table S1 for more details). Changes in abundance of the most important zooplankton species are shown in the barplots for both time points (start and final).

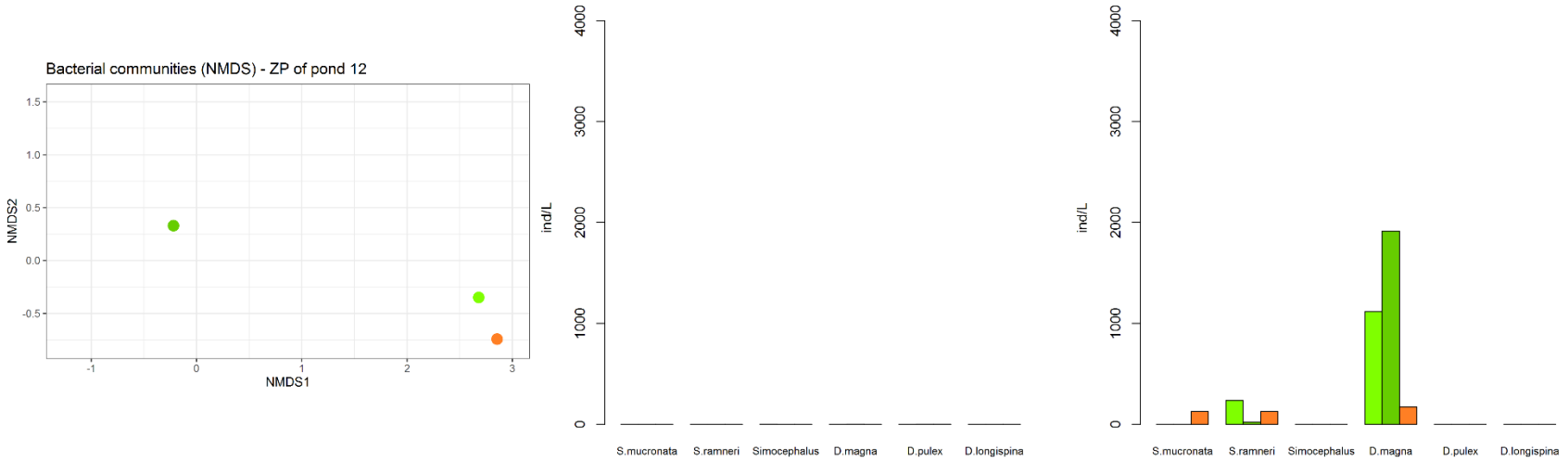

Figure S19. Non-metric multidimensional scaling (NMDS) plot of bacterioplankton communities at the end point of the experiment. These mesocosms were inoculated with zooplankton communities from pond 12 (see Table S1 for more details). Changes in abundance of the most important zooplankton species are shown in the barplots for both time points (start and final).

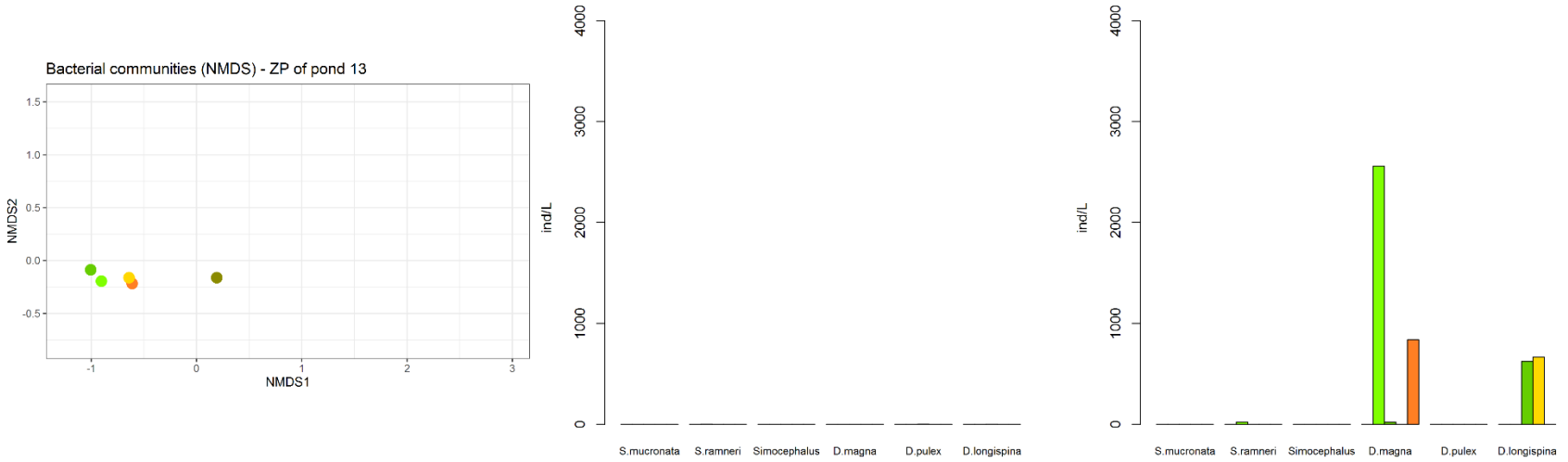

Figure S20. Non-metric multidimensional scaling (NMDS) plot of bacterioplankton communities at the end point of the experiment. These mesocosms were inoculated with zooplankton communities from pond 13 (see Table S1 for more details). Changes in abundance of the most important zooplankton species are shown in the barplots for both time points (start and final).

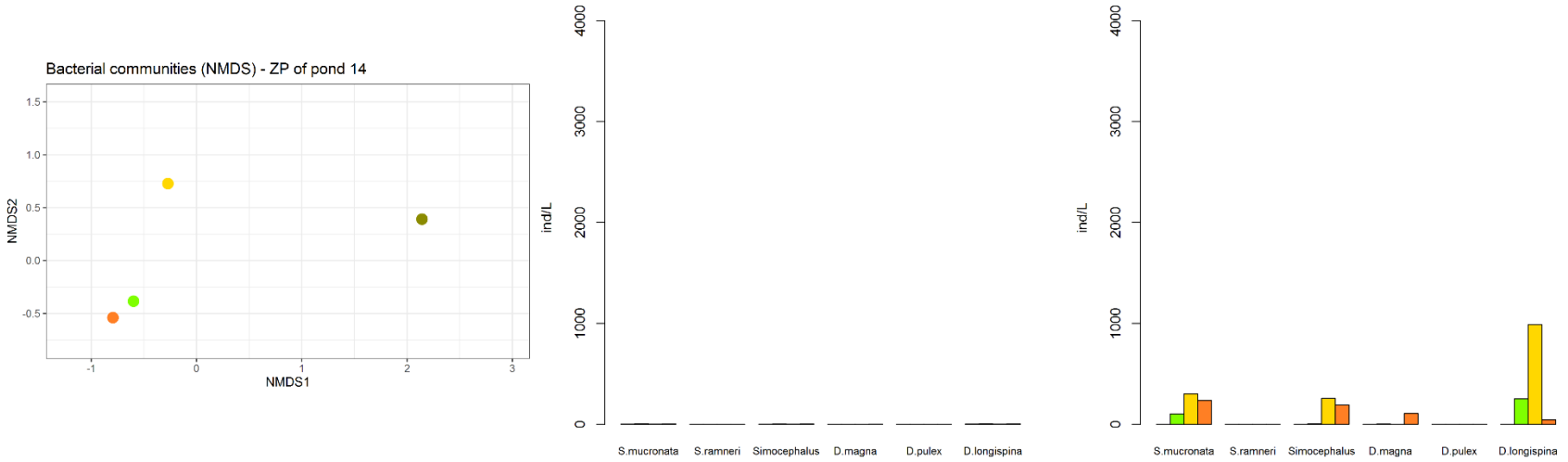

Figure S21. Non-metric multidimensional scaling (NMDS) plot of bacterioplankton communities at the end point of the experiment. These mesocosms were inoculated with zooplankton communities from pond 14 (see Table S1 for more details). Changes in abundance of the most important zooplankton species are shown in the barplots for both time points (start and final).

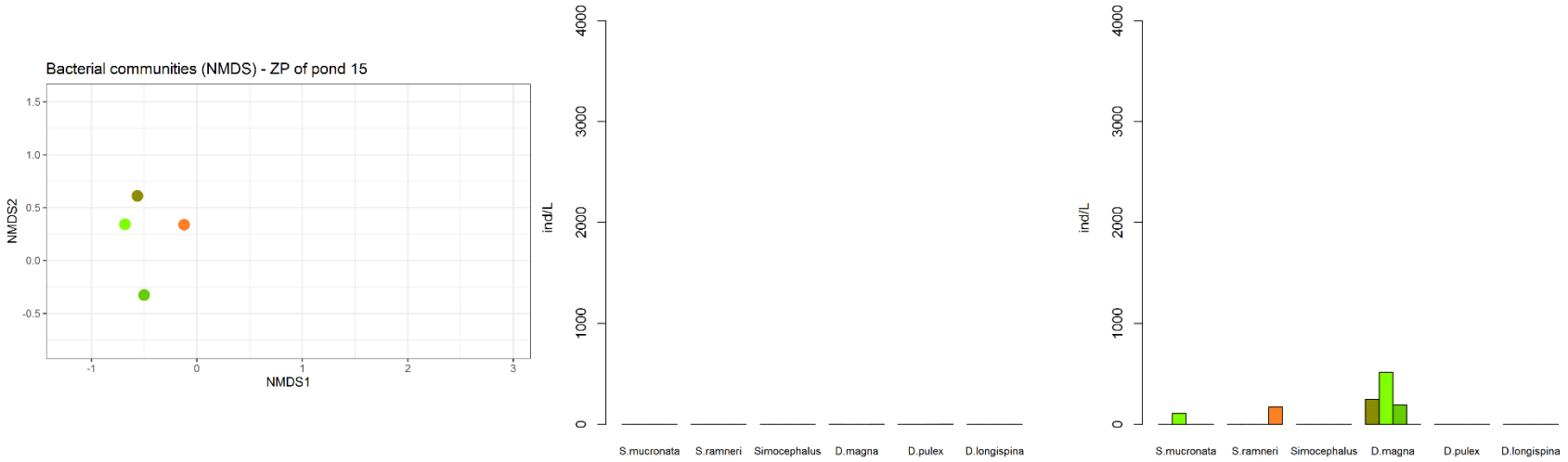

Figure S22. Non-metric multidimensional scaling (NMDS) plot of bacterioplankton communities at the end point of the experiment. These mesocosms were inoculated with zooplankton communities from pond 15 (see Table S1 for more details). Changes in abundance of the most important zooplankton species are shown in the barplots for both time points (start and final).

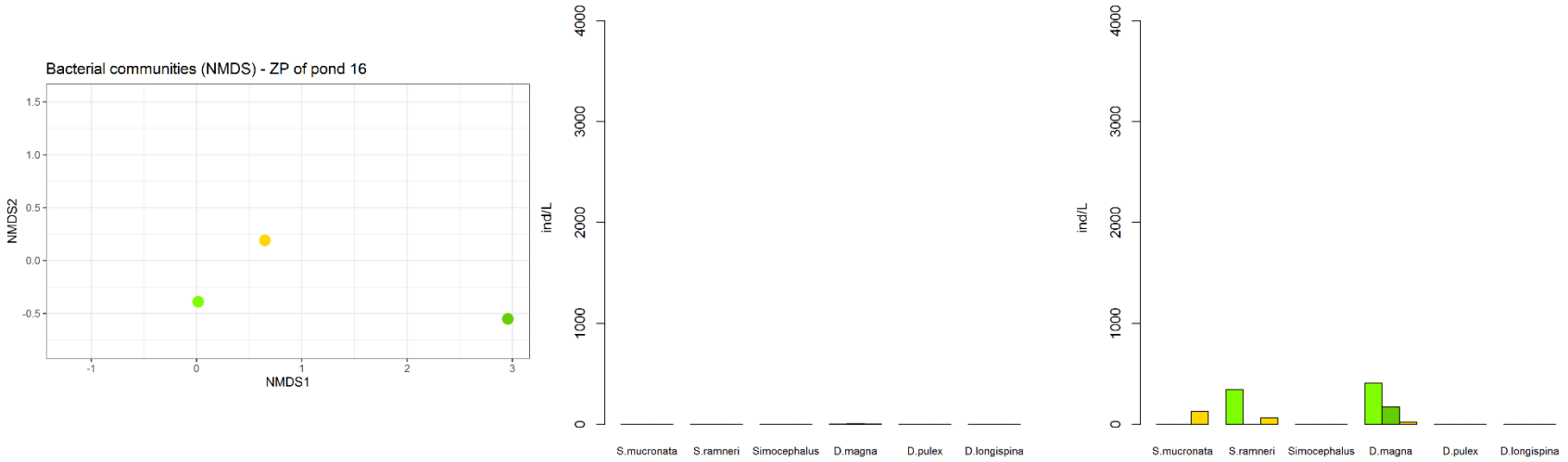

Figure S23. Non-metric multidimensional scaling (NMDS) plot of bacterioplankton communities at the end point of the experiment. These mesocosms were inoculated with zooplankton communities from pond 16 (see Table S1 for more details). Changes in abundance of the most important zooplankton species are shown in the barplots for both time points (start and final).

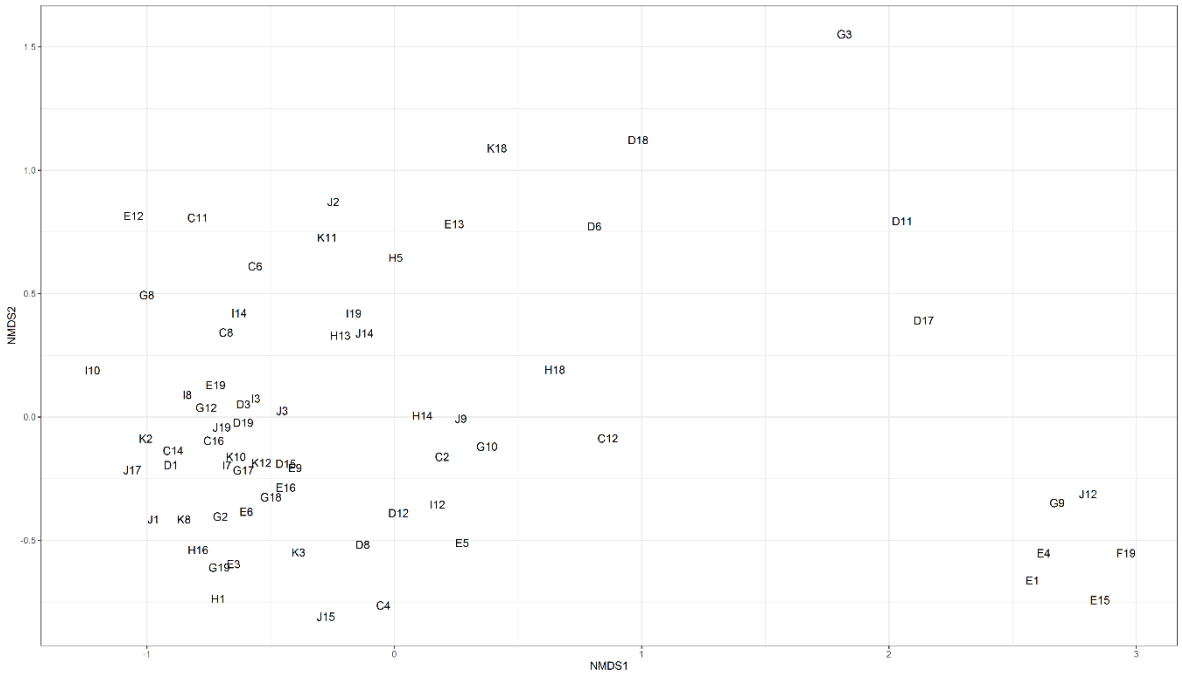

Figure S24. Non-metric multidimensional scaling (NMDS) plot of bacterioplankton communities at the final time point of the experiment. Bacterioplankton communities are displayed by placement codes in the experimental area. Letters represent “rows” (A to K) and numbers, “columns” (1 – 19). See Figure S1 for a visualization of the experimental area.

Chapter VI – General discussion

6.1 General discussion

Bacterioplankton communities play a key role in aquatic food webs and other fundamental biogeochemical cycles (Falkowski et al., 2008; Buchan et al., 2014). Understanding the processes that underly bacterioplankton assembly provides important subsidies on how microbial communities are responding to habitat degradation and the ongoing environmental changes threatening ecosystems and their related functions (Tilman et al., 2017). Community assembly mechanisms are often classified as local and regional processes, where local processes encompass primarily niche mechanisms related to local environmental conditions (more known as species sorting in metacommunity ecology) such as abiotic conditions or biotic interactions (Leibold et al., 2004). On other hand, regional processes are related to the dispersal of organisms among the different local habitats (Langenheder & Lindström, 2019). Microbial dispersal rates have been modelled through vectors as wind, particles, waterways or animals (Mony et al., 2020). The development of high-throughput sequencing methods around the year 2000 made it possible to investigate complete bacterial communities and the mechanisms driving their community structure (Zhou & Ning, 2017). Since then, freshwater bacterioplankton has been widely studied, with species sorting being recognized as the most common assembly mechanism in this system (Van Der Gucht et al., 2007; Souffreau et al., 2015). However, it remained unclear how freshwater bacterioplankton assembly is impacted by anthropogenic pressures such as urbanization and eutrophication, and whether trait-based approaches would enhance our understanding of bacterioplankton assembly. Additionally, biotic factors such as phytoplankton and zooplankton structure have been recognized as important explanatory variables in bacterioplankton assembly studies (Souffreau et al., 2015; Langenheder & Lindström, 2019).

The central aim of this thesis was to understand the drivers of freshwater bacterioplankton assembly along the anthropogenic gradients of urbanization and eutrophication. We were also specifically interested in how important biotic factors such as phytoplankton and zooplankton are for bacterioplankton assembly, and whether trait-based approaches would give us a complementary view on the drivers of bacterioplankton community structure. Using field sampling data of 145 ponds across Belgium and results from a mesocosm experiment manipulating temperature, nutrient status and zooplankton structure, we showed that:

1) bacterioplankton communities are impacted by the anthropogenic gradients of urbanization and eutrophication, although the effects of urbanization were relatively small, and the effects of eutrophication acted mostly indirect through interactions with other trophic levels.
2) on the spatial scale of Belgium local processes were the most important drivers for bacterioplankton community structure compared to regional processes, emphasizing the key importance of species sorting in freshwater bacterioplankton.
3) among the local drivers, biotic factors, and specially zooplankton variables, turned out to be important in explaining the bacterioplankton community structure.
4) trait-based structure was also driven by local environmental factors, but with a different set of variables being important compared to the ones influencing the taxonomy-based structure.

6.2 Impacts of anthropogenic gradients on freshwater bacterioplankton

Humans impact their natural environment drastically, and urbanization, eutrophication and global warming are three important ongoing pressures with huge effects on the biotic and abiotic component of ecosystems (Grimm et al., 2008; Feuchtmayr et al., 2009; Schindler, 2012; Huisman et al., 2018). Freshwater bacterioplankton have been shown to be affected by different environmental changes related to these anthropogenic processes. Diversity loss (Wang et al., 2019), community homogenization (Mykrä et al., 2017) and turnover to specific groups adapted to certain conditions (Newton & McLellan, 2015) are common outcomes related to the responses of freshwater bacterioplankton communities. In this thesis, we combined field surveys along gradients of urbanization (chapter II) and eutrophication (chapter III) in Belgium with a mesocosms experiment manipulating temperature and nutrient status besides zooplankton structure (chapter V). Along the urbanization gradient (chapter II), we did not find direct effects of built-up area, our proxy for urbanization, however, other local conditions (pH, alkalinity, conductivity, total phosphorus, abundance of *Daphnia* and

concentration of copper) were found to be important assembling bacterioplankton communities. pH was found to be partly mediated by urbanization, suggesting that the effects of urbanization on these communities is probably not operating directly, but through changes in the water chemistry. Along the eutrophication gradient (chapter III), suspended matter (a proxy for eutrophication in our study) was shown to be important in driving bacterioplankton community structure, but this effect was mediated by the phytoplankton community. The relation with the phytoplankton community was thus indirect, as the altered phytoplankton structure induced changes in the local environmental conditions under eutrophic conditions. We used suspended matter as proxy for eutrophication rather than nutrient concentrations (e.g., total nitrogen and phosphorus) because it was the variable explaining most part of bacterioplankton community variation. Suspended matter night be a more direct link to the particulate organic matter resulting from the phytoplankton growth, providing resources for the heterotrophic bacteria. Based on the mesocosm experiment (chapter V), no direct effects of warming and eutrophication could be detected, and mainly biotic interactions related to the zooplankton communities were important in driving changes in bacterioplankton structure. These results partially contrast with other studies investigating the effects of urbanization, eutrophication, and warming on freshwater bacterioplankton. We also found turnover for specific groups along the environmental gradients, however the urbanization and eutrophication gradients did not lead to a biotic homogenization of the communities.

In terms of diversity, previous studies have shown mixed results in terms of bacterial diversity being affected by environmental change. For example, we did not find any effects of urbanization on bacterioplankton diversity (chapter II). This pattern was also described for stream bacterial communities sampled along an eutrophication gradient (Simonin et al., 2019). Along the eutrophication gradient, we found a positive relation between concentrations of suspended matter (related to phytoplankton) and diversity of bacterioplankton (chapter III). While an increase (Kiersztyn et al., 2019) and a decrease (Zeng et al., 2019) in lake bacterial diversity along eutrophication gradients were previously described. In the mesocosm experiment, the applied nutrient and temperature treatments did not lead to significant changes in bacterioplankton diversity, however, we found an increase in diversity in relation to the presence of *Daphnia magna*. It thus turns out that, in general, effects of anthropogenic stressors on freshwater bacterioplankton diversity are very divergent and depend on the stressor and other conditions.

Our studies were to a certain extent limited in covering the selected environmental gradients. The urbanization gradient in Belgium did not include pristine areas, with most part of the sampled areas being affected by human activities. Although we tried to evaluate the direct effects of anthropogenic stressors, for example, using values of built-up area as proxy for urbanization, the effects on aquatic community structure were predominantly detectable via changes in the local environmental conditions. We are aware that factorial studies have advantages to compare environmental conditions that can highlight more extreme differences and usually are easier to analyse (e.g., disturbed x non-disturbed) (Mykrä et al., 2017). However, we think our approach of using environmental gradients represents in a better way the variability found in the natural world, across spatial and temporal scales. In our studies, we observed more gradual community changes along the axes of variation by using environmental gradients (chapter II and III). Our experiment, in terms of environmental variables had much higher eutrophication values compared to the ones found in shallow lakes in the observational samples (chapter III). However, this was not translated to an increase in phytoplankton growth in all nutrient-enriched mesocosms. Even in more controlled conditions, our experiment had limitations in representing the different scenarios, probably due to the extra levels of complexity (three nutrient-enrichment levels, two temperatures and 16 different zooplankton communities). But the hypertrophic conditions found in our experiment are not that uncommon in the real world, as illustrated by the current increase in algae blooms across the planet (Huisman et al., 2018).

6.3 Drivers of freshwater bacterioplankton metacommunities

Multiple studies have investigated the drivers of freshwater bacterioplankton, often within a metacommunity perspective (Adams et al., 2014; Roguet et al., 2015; Huber et al., 2020). In this thesis, we aimed to study the drivers of freshwater bacterioplankton along specific anthropogenic gradients, being urbanization (chapter II) and eutrophication (chapter III), as it has been shown that environmental change might alter the relative importance of community assembly processes (Ren et al., 2015). We also incorporated biotic variables related to zooplankton and phytoplankton as explanatory variables (chapter II, III, IV and V) as these were often disregarded in previous metacommunity studies. Lastly, we also studied whether trait-based approaches could improve our understanding of bacterioplankton assembly

(chapter IV). The main results of these different components will be discussed in the following paragraphs.

Species sorting: a key process for bacterioplankton assembly

Metacommunity studies have repeatedly found species sorting as one of the most important mechanism driving changes on bacterioplankton assembly across distinct systems – marine, freshwater – and along different environmental gradients and approaches – observational and experimental (Van Der Gucht et al., 2007; Yeh et al., 2015). Even with several studies in the past decade showing that other processes can also be of importance (e.g., dispersal limitation, mass effects and neutrality) (Bell, 2010; Souffreau et al., 2014; Roguet et al., 2015), a part of the variation in bacterioplankton's distribution is always driven by environmental conditions at the local scale. In this thesis as well, environmental selection, or species sorting in a metacommunity context, has been the main process driving bacterioplankton assembly, even considering the two different environmental gradients studied (urbanization and eutrophication), and the two approaches used (field observation and experimental mesocosms).

One of the reasons why environmental selection is often detected as the most important process for bacterioplankton assembly are the limitations to assess regional factors. Generally, local environmental variables are easier to be measured and are therefore the most common factors included as explanatory variables in ecological studies, including microbial studies (Lindström & Langenheder, 2012). We used indirect methods to estimate spatial factors (e.g., geographical coordinates for the linear components and MEMs – Moran's eigenvector maps for the non-linear spatial components). Although we still used more traditional methods, we tried to improve the accuracy of the spatial component, using a hierarchical spatial design in chapter II to disentangle the effects of urbanization across different spatial scales. For microbial communities, only a small proportion of variation (2%) was explained by urbanization at the spatial scale of 200 m around the shallow lakes.

In our studies, larger amounts of variation in community composition could not be explained either by local or regional factors. This is a general phenomenon in microbial metacommunity studies (Zhou & Ning, 2017). This pattern could be related to the fact that we did not measure the factors that are driving changes in community structure. We are also aware of the limitations of the variation partitioning method in explaining community

variation (Gilbert & Bennett, 2010). Variation partitioning has been criticized in failing to represent local and regional processes and to produce inflated values of explained variation. However, we carefully interpreted our results in terms of the explained variation generated by the variation partitioning models. With the results and possible relationships found here, new hypotheses can be derived and then, more specific statistical methods can be used to address it.

pH as an important abiotic driver for bacterioplankton community structure

Among the variables selected driving bacterioplankton structure, pH was found to be an important variable in the two different environmental contexts (urbanization and eutrophication) (chapter II and III). pH is known to strongly influence microbial distribution in different environments such as soil (Xu et al., 2014) and freshwater systems (Ren et al., 2015). Since most microorganisms have an intracellular pH within one pH unit of neutral range, this restricts their ability to survive or reproduce under certain pH conditions. pH therefore has a direct effect on bacterioplankton community structure (Newton et al., 2011). Additionally, pH can also indirectly induce changes in bacterial composition such as through nutrient availability (Kampe et al., 2010) and cationic metal solubility (Lu et al., 2010). Although we did not have extreme values of pH (acid or alkaline), it explained part of the bacterioplankton variation over the urbanization and eutrophication gradients. In chapter IV, pH was selected influencing the taxonomic-based structure, but not the trait-based structure. pH was also selected as an important variable structuring bacterioplankton in the mesocosm experiment (chapter V), with an increase in pH values along the 12 weeks. Specific pH conditions are known to select specific microbial species and groups (e.g., Acidobacteria), because specific traits are needed to survive in these conditions. These traits are often phylogenetically conserved, and they are often important to select which species can survive in specific conditions (Martiny et al., 2015).

The importance of biotic interactions on bacterioplankton community structure

Among the different environmental factors measured in this thesis to evaluate the role of local environmental factors on bacterioplankton assembly, biotic interactions were found directly and indirectly being relevant under the different scenarios studied (chapters II, III, IV and V). Biotic interactions have been included as explanatory variables in microbial studies more

recently (Verreydt et al., 2012; Goberna et al., 2014; Souffreau et al., 2015), and it has been shown that these interactions could be important in structuring bacterioplankton through predation (Chow et al., 2014) and host-associated microbiota (Qi et al., 2009). Biotic interactions could also be the result of indirect biotic effects as observed in chapter III, where bacterioplankton was indirectly structured by the phytoplankton community through the concentration of suspended matter. Because our studies are only observational, we cannot disentangle the mechanisms that are at the base of these indirect interactions. However, our studies reinforce the importance of including biotic interactions as explanatory variables in bacterioplankton metacommunity analyses.

6.4 Trait-based structure as a complement in the study of bacterioplankton communities

Trait-based approaches have been put forward in metacommunity analyses because traits might be more meaningful than taxon-based approaches since traits might be more related with specific niches and functional roles in the ecosystems (Martiny et al., 2015; Yang, 2021). Trait-based approaches could also be more informative to track environmental changes related to anthropogenic pressures. In this thesis, two different trait-based methods were used to evaluate whether traits can help us to fully understand the drivers of bacterioplankton assembly (chapter IV). We found that functional group annotations were driven by a partly different set of local variables (suspended matter concentration, pH, abundance of *Daphnia*, concentrations of available and total phosphorus) compared to the taxonomical structure (concentrations of dissolved oxygen, total phosphorus and suspended matter). This result is in line with previous studies where this pattern was also described (Bissett et al., 2010; Ruiz-González et al., 2015b), and indicates that trait-based approaches offer indeed a complementary view on the drivers structuring bacterioplankton structure.

While it is not possible to compare the amount of explained community variation among different datasets (in our case trait-based and taxon-based), our results indicate that local environmental conditions were more important than regional factors for both dataset types. Comparing the most important variables explaining the taxonomic-based structure in chapter IV (suspended matter concentration, pH, abundance of *Daphnia*, concentrations of

available and total phosphorus), we can see an overlap among the variables found to be significant in chapter II and III. The set of ponds and shallow lakes used in chapter IV is not a complete combination of the set of ponds used in chapter II and III, as only a subset of these ponds was analysed in chapter IV. This strengthens the idea that these variables are indeed important for the bacterioplankton taxonomic structure of shallow ponds in Belgium, and that the trait-based approach complements our understanding of its structure. For the trait-based community structure, abiotic variables such as dissolved oxygen and available and total phosphorus were selected, variables that are more related to the processing of organic matter and other ecological functions that are known to be intermediated by bacteria in aquatic systems (Briand et al., 2004; Smith & Prairie, 2004). Therefore, the use of trait-based approaches shows to be a valuable approach to improve the study of bacterioplankton assembly, providing a more complete picture of the drivers that are important.

Although local environmental drivers apparently can explain part of the variation in the trait-based structure for functional group annotations, no variables were selected for the carbon consumption profiles. It is important to note that specific variables (e.g., quality of organic matter) might be key to understand patterns of carbon consumption in freshwater bodies (Ruiz-González et al., 2015b). We therefore suggest that future studies that studying carbon utilization of bacterioplankton communities should also measure organic matter variables.

6.5 Sequencing methods and statistical consequences

In this thesis, two different sequencing methods have been used: 454 pyrosequencing and Illumina MiSeq. 454 pyrosequencing tends to provide a lower number of sequences, although with longer reads compared to Illumina MiSeq (Pollock et al., 2018). We were aware of this shortcoming, so we standardized our samples to a minimum of 1000 sequences per sample when 454 pyrosequencing was used. This number of sequences is considered sufficient to described trends and patterns of beta-diversity (Lundin et al., 2012). In chapter II, a lower number of sequences (688 per sample) was used for standardization to increase the number of samples, but rarefaction curves were plotted to check for species saturation before further analyses. Besides, all statistical analyses were also carried out with 1000 sequences per sample and led to the same patterns. However, alpha diversity is much affected by low number of

sequences per sample, and we therefore limited our comparisons relatively to the other samples within our metacommunity. Additionally, we also used the Shannon-Wiener index as diversity measure next to richness to try to overcome the sequencing depth issue. This index also considers on how even the communities are. Since in our studies we only detect the most common OTUs with the sequencing depth used, the Shannon-Wiener index complements the information about the diversity of these communities.

Another general consequence of high throughput sequencing methods was the discovery that almost the totality of microbial communities is highly diverse, but with most part of the species being rare (Pedrós-Alió, 2012). This pattern brings challenges to the analyses on how to deal with the high number of species showing low abundances and frequency. We tried to overcome this issue with different approaches to improve our comparisons and statistical analyses. In chapter II, community data were analysed following the idea that most part of the community variation is driven by the most abundant/common species, so we ran our analyses with different cut-off levels of relative abundance (≥ 0.1% and ≥ 1%). Since most part of the community variation is related to the most abundant/common species, using this approach we were able to capture the most important patterns in community structure related to urbanization. In chapter III, we used PCoA (Principal Coordinate Analysis) eigenvectors to synthetize community variation allowing comparisons among the other two trophic levels (phytoplankton and zooplankton communities), since in number of response variables (number of OTUs, number of phytoplankton and zooplankton species) each dataset could differ in two orders of magnitude. In chapter IV, we used the same approach as in chapter III, but this time to compare the variation among taxonomic, functional group annotations and carbon utilization profiles (number of OTUs – taxonomy-based; number of functional profiles – trait-based; number of carbon sources consumed – trait-based). In chapter V, the co-occurrence analyses were conducted only with the most abundant OTUs across mesocosms to reduce the number of spurious correlations or correlations between very rare species across samples. In general, there was the need to choose different approaches in each study to reduce the possible noise caused by the inclusion of rare species. It was not our aim to study in detail how these rare species were driven by local and spatial factors. Rare species have been the focus of other studies evaluating for example whether different variables are important for the community structure of rare species compared to the abundant ones (Baltar et al., 2015; Liu et al., 2015). Due to their low abundance, rare species

are more prone to be driven by regional processes, including ecological drift (Zhou & Ning, 2017). Our approaches are designed for the interpretation of more general patterns in bacterioplankton assembly, linked to the most abundant species/OTUs.

An additional consequence of the shortcomings related to our methodological choices is the lack of resolution for the classification of taxa for example by FAPROTAX. This might be overcome by a more robust taxon identification that could improve the functional assignments. A classification based on exact sequence variants (more known as amplicon sequence variant – ASV) could improve the functional assignments since ASVs rely on exact nucleotide sequences of marker genes, rather than a similarity percentage (Callahan et al., 2017).

6.6 Perspectives and future work

For the past decade, many advances were made in understanding the distribution of microbial communities. It is now well accepted that microbes are not “everywhere” and that indeed other factors play a role in their distribution across space and not only local environmental conditions (Martiny et al., 2006). However, it is clear from the results of this thesis that species sorting remains a preponderant factor for bacterioplankton assembly. The use of environmental gradients could be extended to evaluate the role of species sorting in different contexts such as along a pH gradient that proved to be a determinant factor driving bacterioplankton structure. Additionally, biotic interactions must be included as potential drivers in future studies that are wishing to quantify the importance of local factors in bacterioplankton assembly. Groups recognized to drive changes in bacterioplankton structure such as viruses (through viral lysis) and heterotrophic nanoflagellates (through direct predation) should be included as explanatory variables in future studies of bacterioplankton assembly. The recent understanding of the role of the microbiome of different species on ecological processes opened a new research avenue that could bring light to the interactions with other organisms influencing bacterioplankton assembly (Qi et al., 2009; Sison-Mangus et al., 2015; Knight et al., 2018). A recurring issue with studies that try to disentangle the relative importance of local and regional drivers is that local factors are easier to measure and easier to estimate compared to regional ones (Lindström & Langenheder, 2012). New statistical and experimental approaches have been used to better understand the role of regional processes

on microbial distribution. Null model approaches (Stegen et al., 2012; Zhou & Ning, 2017) and Hierarchical Modelling of Species Communities (HMSC) (Ovaskainen et al., 2017) are examples of statistical tools that could complement the statistical approaches used in this thesis. Experiments have been used to test whether regional processes are important for bacterioplankton assembly manipulating dispersal rates (Souffreau et al., 2014). Future studies could combine the experimental setup that we used (stressors + zooplankton richness gradient) with manipulation of bacterioplankton dispersal rates among mesocosms to evaluate the effects of controlled dispersal on the biotic interactions between bacterioplankton and zooplankton communities.

It is important to notice that the studies in this PhD thesis only focussed on how the spatial component influences bacterioplankton assembly, although also the incorporation of the temporal scale is fundamental to better understand metacommunity assembly. The order and the time in which species arrive from the regional pool can influence the species abundances within the metacommunity (Fukami, 2015). Also, it has been shown that past environmental conditions could be a better predictor of actual bacterioplankton community structure. Andersson et al. (2014) studying bacterioplankton of rock pools was able to demonstrate the significant effect of environmental conditions measured 8 days earlier than the actual community sampling. Therefore, future studies could improve their capacity to explain variation in community composition if they incorporate the temporal component.

In terms of the trait-based structure, here we used indirect approaches such as the categorization by functional annotations where one evaluates potential functions related to different taxa, and the carbon consumption profiles by which the ecological function is measured for the whole community. However, to implement microbial traits as biomonitoring tools to track environmental change, it has been proposed that future studies should focus on the dominant taxa that are responsible for most part of the ecosystem services provided by a community (Yang, 2021). Additionally, data based on the analysis of the functional gene composition through metagenomics will provide a more complete picture on how bacteria in freshwater systems are responding to the impacts of the environmental change beyond taxonomic information.

References

Adams, H. E., B. C. Crump, & G. W. Kling, 2014. Metacommunity dynamics of bacteria in an arctic lake: The impact of species sorting and mass effects on bacterial production and biogeography. Frontiers in Microbiology 5: 82.

Afshinnekoo, E., C. Meydan, S. Chowdhury, D. Jaroudi, C. Boyer, N. Bernstein, J. M. Maritz, D. Reeves, J. Gandara, S. Chhangawala, S. Ahsanuddin, A. Simmons, T. Nessel, B. Sundaresh, E. Pereira, E. Jorgensen, S. O. Kolokotronis, N. Kirchberger, I. Garcia, D. Gandara, S. Dhanraj, T. Nawrin, Y. Saletore, N. Alexander, P. Vijay, E. M. Hénaff, P. Zumbo, M. Walsh, G. D. O'Mullan, S. Tighe, J. T. Dudley, A. Dunaif, S. Ennis, E. O'Halloran, T. R. Magalhaes, B. Boone, A. L. Jones, T. R. Muth, K. S. Paolantonio, E. Alter, E. E. Schadt, J. Garbarino, R. J. Prill, J. M. Carlton, S. Levy, & C. E. Mason, 2015. Geospatial Resolution of Human and Bacterial Diversity with City-Scale Metagenomics. Cell Systems 1: 72–87.

Alberti, M., C. Correa, J. M. Marzluff, A. P. Hendry, E. P. Palkovacs, K. M. Gotanda, V. M. Hunt, T. M. Apgar, & Y. Zhou, 2017. Global urban signatures of phenotypic change in animal and plant populations. Proceedings of the National Academy of Sciences of the United States of America 114: 8951–8956.

Alberti, M., J. M. Marzluff, E. Shulenberger, G. Bradley, C. Ryan, & C. Zumbrunnen, 2003. Integrating human into ecology: Opportunities and challenges for studying urban ecosytems. BioScience 53: 1169–1179.

Anderson, M. J., 1999. Distinguishing direct from indirect effects of grazers in intertidal estuarine assemblages. Journal of Experimental Marine Biology and Ecology 234: 199–218.

Anderson, M. J., 2001. A new method for non parametric multivariate analysis of variance. Austral Ecology 26: 32–46.

Anderson, M. J., 2006. Distance-based tests for homogeneity of multivariate dispersions. Biometrics 62: 245–253.

Anderson, M. J., & T. J. Willis, 2003. Canonical analysis of principal coordinates: A useful method of constrained ordination for ecology. Ecology 84: 511–525.

Andersson, M. G. I., M. Berga, E. S. Lindström, & S. Langenheder, 2014. The spatial structure of bacterial communities is influenced by historical environmental conditions. Ecology 95: 1134–1140.

Bai, Y., Q. Shi, D. Wen, Z. Li, W. a Jefferson, C. Feng, & X. Tang, 2012. Bacterial communities

in the sediments of Dianchi lake, a partitioned eutrophic waterbody in China. PLoS ONE 7: e37796.

Baltar, F., J. Palovaara, M. Vila-Costa, G. Salazar, E. Calvo, C. Pelejero, C. Marrasé, J. M. Gasol, & J. Pinhassil, 2015. Response of rare, common and abundant bacterioplankton to anthropogenic perturbations in a Mediterranean coastal site. FEMS Microbiology Ecology 91(6): fiv058.

Barberán, A., A. Fernández-Guerra, B. J. M. Bohannan, & E. O. Casamayor, 2012. Exploration of community traits as ecological markers in microbial metagenomes. Molecular Ecology 21: 1909–1917.

Barnosky, A. D., N. Matzke, S. Tomiya, G. O. U. Wogan, B. Swartz, T. B. Quental, C. Marshall, J. L. McGuire, E. L. Lindsey, K. C. Maguire, B. Mersey, & E. A. Ferrer, 2011. Has the Earth's sixth mass extinction already arrived? Nature 471: 51–57.

Baselga, A., & C. D. L. Orme, 2012. Betapart: An R package for the study of beta diversity. Methods in Ecology and Evolution 3: 808–812.

Bastian, M., S. Heymann, & M. Jacomy, 2009. Gephi: an open source software for exploring and manipulating networks. International AAAI Conference on Weblogs and Social Media.

Becker, R. A., A. R. Wilks, R. Brownrigg, T. P. Minka, & A. Deckmyn, 2018. maps: Draw Geographical Maps. https://cran.r-project.org/package=maps.

Beisner, B. E., P. R. Peres-Neto, E. S. Lindström, A. Barnett, & M. L. Longhi, 2006. The role of environmental and spatial processes in structuring lake communities from bacteria to fish. Ecology 87: 2985–2991.

Beklioğlu, M., M. Meerhoff, T. A. Davidson, K. A. Ger, K. Havens, & B. Moss, 2016. Preface: Shallow lakes in a fast changing world. Hydrobiologia 778: 9–11.

Bell, T., 2010. Experimental tests of the bacterial distance-decay relationship. ISME Journal 4: 1357–1365.

Beneš, J., P. Kepka, & M. Konvička, 2003. Limestone Quarries as Refuges for European Xerophilous Butterflies. Conservation Biology 17: 1058–1069.

Berga, M., Ö. Östman, E. S. Lindström, & S. Langenheder, 2015. Combined effects of zooplankton grazing and dispersal on the diversity and assembly mechanisms of bacterial metacommunities. Environmental Microbiology 17: 2275–2287.

Birtel, J., & B. Matthews, 2016. Grazers structure the bacterial and algal diversity of aquatic metacommunities. Ecology 97: 3472–3484.

Bissett, A., A. E. Richardson, G. Baker, S. Wakelin, & P. H. Thrall, 2010. Life history determines biogeographical patterns of soil bacterial communities over multiple spatial scales. Molecular Ecology 19: 4315–4327.

Blanchet, F. G., P. Legendre, & D. Borcard, 2008. Forward selection of explanatory variables. Ecology 89: 2623–2632.

Böer, S. I., S. I. C. Hedtkamp, J. E. E. Van Beusekom, J. A. Fuhrman, A. Boetius, & A. Ramette, 2009. Time- and sediment depth-related variations in bacterial diversity and community structure in subtidal sands. ISME Journal 3: 780–791.

Borcard, D., F. Gillet, & P. Legendre, 2011. Numerical Ecology with R. Numerical Ecology with R. Springer, New York.

Bowman, J. S., & H. W. Ducklow, 2015. Microbial communities can be described by metabolic structure: A general framework and application to a seasonally variable, depth-stratified microbial community from the coastal West Antarctic Peninsula. PLoS ONE 10: 1–18.

Brans, K. I., J. M. T. Engelen, C. Souffreau, & L. De Meester, 2018. Urban hot-tubs: Local urbanization has profound effects on average and extreme temperatures in ponds. Landscape and Urban Planning 176: 22–29.

Briand, E., O. Pringault, S. Jacquet, & J. P. Torréton, 2004. The use of oxygen microprobes to measure bacterial respiration for determining bacterioplankton growth efficiency. Limnology and Oceanography: Methods 2: 406–416.

Brum, J. R., B. L. Hurwitz, O. Schofield, H. W. Ducklow, & M. B. Sullivan, 2016. Seasonal time bombs: Dominant temperate viruses affect Southern Ocean microbial dynamics. ISME Journal 10: 437–449.

Buchan, A., G. R. LeCleir, C. A. Gulvik, & J. M. González, 2014. Master recyclers: features and functions of bacteria associated with phytoplankton blooms. Nature Reviews Microbiology 12: 686–698.

Burns, A. R., W. Z. Stephens, K. Stagaman, S. Wong, J. F. Rawls, K. Guillemin, & B. J. M. Bohannan, 2016. Contribution of neutral processes to the assembly of gut microbial communities in the zebrafish over host development. ISME Journal 10: 655–664.

Callahan, B. J., P. J. McMurdie, & S. P. Holmes, 2017. Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. ISME Journal 11: 2639–2643.

Callens, M., E. Macke, K. Muylaert, P. Bossier, B. Lievens, M. Waud, & E. Decaestecker, 2016. Food availability affects the strength of mutualistic host-microbiota interactions in *Daphnia magna*. ISME Journal 10: 911–920.

Cañedo-Argüelles, M., C. P. Hawkins, B. J. Kefford, R. B. Schafer, B. J. Dyack, S. Brucet, D. Buchwalter, J. Dunlop, O. Fror, J. Lazorchak, E. Coring, H. R. Fernandez, W. Goodfellow, A. L. G. Achem, S. Hatfield-Dodds, B. K. Karimov, P. Mensah, J. R. Olson, C. Piscart, N. Prat, S. Ponsa, C.-J. Schulz, & A. J. Timpano, 2016. Saving freshwater from salts. Science 351: 914–916.

Caruso, T., Y. Chan, D. C. Lacap, M. C. Y. Lau, C. P. McKay, & S. B. Pointing, 2011. Stochastic and deterministic processes interact in the assembly of desert microbial communities on a global scale. ISME Journal 5: 1406–1413.

Chadwick, M. A., D. R. Dobberfuhl, A. C. Benke, A. D. Huryn, K. Suberkropp, & J. E. Thiele, 2006. Urbanization affects stream ecosystem function by altering hydrology, chemistry, and biotic richness. Ecological Applications 16: 1796–1807.

Chase, J. M., 2003. Community assembly: When should history matter?. Oecologia 136: 489–498.

Chase, J. M., 2010. Stochastic community assembly causes higher biodiversity in more productive environments. Science 328: 1388–1391.

Chase, J. M., E. G. Biro, W. A. Ryberg, & K. G. Smith, 2009. Predators temper the relative importance of stochastic processes in the assembly of prey metacommunities. Ecology Letters 12: 1210–1218.

Chen, J., P. Wang, C. Wang, X. Wang, L. Miao, S. Liu, Q. Yuan, & S. Sun, 2019. Fungal community demonstrates stronger dispersal limitation and less network connectivity than bacterial community in sediments along a large river. Environmental Microbiology 22(3): 832-849.

Chen, W., Y. Pan, L. Yu, J. Yang, & W. Zhang, 2017. Patterns and processes in marine microeukaryotic community biogeography from Xiamen coastal waters and intertidal sediments, southeast China. Frontiers in Microbiology 8: 1912.

Chow, C. E. T., D. Y. Kim, R. Sachdeva, D. A. Caron, & J. A. Fuhrman, 2014. Top-down

controls on bacterial community structure: Microbial network analysis of bacteria, T4-like viruses and protists. ISME Journal 8: 816–829.

Christoffersen, K., N. Andersen, M. Søndergaard, L. Liboriussen, & E. Jeppesen, 2006. Implications of climate-enforced temperature increases on freshwater pico- and nanoplankton populations studied in artificial ponds during 16 months. Hydrobiologia 560: 259–266.

Comte, J., M. Berga, I. Severin, J. B. Logue, & E. S. Lindström, 2017. Contribution of different bacterial dispersal sources to lakes: Population and community effects in different seasons. Environmental Microbiology 19: 2391–2404.

Comte, J., & P. A. Del Giorgio, 2010. Linking the patterns of change in composition and function in bacterioplankton successions along environmental gradients. Ecology 91: 1466–1476.

Concepción, E. D., M. Moretti, F. Altermatt, M. P. Nobis, & M. K. Obrist, 2015. Impacts of urbanisation on biodiversity: The role of species mobility, degree of specialisation and spatial scale. Oikos 124: 1571–1582.

Conway, T. M., 2007. Impervious surface as an indicator of pH and specific conductance in the urbanizing coastal zone of New Jersey, USA. Journal of Environmental Management 85: 308–316.

Cottenie, K., 2005. Integrating environmental and spatial processes in ecological community dynamics. Ecology Letters 8: 1175–1182.

Crevecoeur, S., C. Ruiz-González, Y. T. Prairie, & P. A. Del Giorgio, 2019. Large-scale biogeography and environmental regulation of methanotrophic bacteria across boreal inland waters. Molecular Ecology 28: 4181–4196.

Dai, W., J. Zhang, Q. Tu, Y. Deng, Q. Qiu, & J. Xiong, 2017. Bacterioplankton assembly and interspecies interaction indicating increasing coastal eutrophication. Chemosphere 177: 317–325.

Darimont, C. T., S. M. Carlson, M. T. Kinnison, P. C. Paquet, T. E. Reimchen, & C. C. Wilmers, 2009. Human predators outpace other agents of trait change in the wild. Proceedings of the National Academy of Sciences of the United States of America 106: 952–954.

Davies, P. J., I. A. Wright, O. J. Jonasson, & S. J. Findlay, 2010. Impact of concrete and PVC pipes on urban water chemistry. Urban Water Journal 7: 233–241.

De Bie, T., L. De Meester, L. Brendonck, K. Martens, B. Goddeeris, D. Ercken, H. Hampel, L. Denys, L. Vanhecke, K. Van der Gucht, J. Van Wichelen, W. Vyverman, & S. A. J. Declerck, 2012. Body size and dispersal mode as key traits determining metacommunity structure of aquatic organisms. Ecology Letters 15: 740–747.

De Meester, L., S. Declerck, R. Stoks, G. Louette, F. Van De Meutter, T. De Bie, E. Michels, & L. Brendonck, 2005. Ponds and pools as model systems in conservation biology, ecology and evolutionary biology. Aquatic Conservation: Marine and Freshwater Ecosystems 15: 715–725.

De Meester, L., & J. Pantel, 2014. Eco-evolutionary dynamics in freshwater systems. Journal of Limnology 73: 193–200.

De Wit, R., & T. Bouvier, 2006. "Everything is everywhere, but, the environment selects"; what did Baas Becking and Beijerinck really say? Environmental Microbiology 8: 755–758.

Declerck, S., T. De Bie, D. Ercken, H. Hampel, S. Schrijvers, J. Van Wichelen, V. Gillard, R. Mandiki, B. Losson, D. Bauwens, S. Keijers, W. Vyverman, B. Goddeeris, L. De Meester, L. Brendonck, & K. Martens, 2006. Ecological characteristics of small farmland ponds: Associations with land use practices at multiple spatial scales. Biological Conservation 131: 523–532.

Degans, H., & L. De Meester, 2002. Top-down control of natural phyto- and bacterioplankton prey communities by *Daphnia magna* and by the natural zooplankton community of the hypertrophic Lake Blankaart. Hydrobiologia 479: 39–49.

Dickerson, T. L., & H. N. Williams, 2014. Functional Diversity of Bacterioplankton in Three North Florida Freshwater Lakes over an Annual Cycle. Microbial Ecology 67: 34–44.

Downing, J. A., 2010. Emerging global role of small lakes and ponds: little things mean a lot. Limnetica 29: 9–24.

Drakare, S., & A. Liess, 2010. Local factors control the community composition of cyanobacteria in lakes while heterotrophic bacteria follow a neutral model. Freshwater Biology 55: 2447–2457.

Dray, A. S., G. Blanchet, D. Borcard, G. Guenard, T. Jombart, G. Larocque, P. Legendre, N. Madi, & H. H. Wagner, 2016. adespatial: Multivariate Multiscale Spatial Analysis. R package, Version: 0.0-7, https://cran.r-project.org/package=adespatial.

Dray, S., P. Legendre, & P. R. Peres-Neto, 2006. Spatial modelling: a comprehensive framework for principal coordinate analysis of neighbour matrices (PCNM). Ecological

Modelling 196: 483–493.

Duarte, L. D. S., P. V. Prieto, & V. D. Pillar, 2012. Assessing spatial and environmental drivers of phylogenetic structure in Brazilian Araucaria forests. Ecography 35: 952–960.

Dusa, A., 2021. venn: Draw Venn Diagrams. R package, Version 1.11, https://github.com /dusadrian/venn.

Eckert, E. M., & J. Pernthaler, 2014. Bacterial epibionts of *Daphnia*: A potential route for the transfer of dissolved organic carbon in freshwater food webs. ISME Journal 8: 1808–1819.

Eiler, A., & S. Bertilsson, 2004. Composition of freshwater bacterial communities associated with cyanobacterial blooms in four Swedish lakes. Environmental Microbiology 6: 1228–1243.

Epskamp, S., & S. Stuber, 2017. semPlot: Path Diagrams and Visual Analysis of Various SEM Packages' Output. R package, https://cran.r-project.org/package=semPlot.

Escalas, A., L. Hale, J. W. Voordeckers, Y. Yang, M. K. Firestone, L. Alvarez-Cohen, & J. Zhou, 2019. Microbial functional diversity: From concepts to applications. Ecology and Evolution 9: 12000–12016.

Falkowski, P. G., T. Fenchel, & E. F. Delong, 2008. The microbial engines that drive earth's biogeochemical cycles. Science 320: 1034–1039.

Feuchtmayr, H., R. Moran, K. Hatton, L. Connor, T. Heyes, B. Moss, I. Harvey, & D. Atkinson, 2009. Global warming and eutrophication: Effects on water chemistry and autotrophic communities in experimental hypertrophic shallow lake mesocosms. Journal of Applied Ecology 46: 713–723.

Fierer, N., 2017. Embracing the unknown: Disentangling the complexities of the soil microbiome. Nature Reviews Microbiology 15: 579–590.

Fisher, J. C., R. J. Newton, D. K. Dila, & S. L. McLellan, 2015. Urban microbial ecology of a freshwater estuary of Lake Michigan. Elementa 3: 000064.

Flössner, D., 2000. Die Haplopoda und Cladocera (ohne Bosminidae) Mitteleuropas. Backhuys, Leiden, Germany.

Frank, K. T., B. Petrie, J. S. Choi, & W. C. Leggett, 2005. Trophic Cascades in a Formerly Cod-Dominated Ecosystem. Science 1621–1624.

Fukami, T., 2004. Assembly history interacts with ecosystem size to influence species diversity. Ecology 85: 3234–3242.

Fukami, T., 2015. Historical Contingency in Community Assembly: Integrating Niches, Species Pools, and Priority Effects. Annual Review of Ecology, Evolution, and Systematics 46: 1–23.

Fukami, T., H. J. E. Beaumont, X. X. Zhang, & P. B. Rainey, 2007. Immigration history controls diversification in experimental adaptive radiation. Nature 446: 436–439.

Fukami, T., I. A. Dickie, J. Paula Wilkie, B. C. Paulus, D. Park, A. Roberts, P. K. Buchanan, & R. B. Allen, 2010. Assembly history dictates ecosystem functioning: Evidence from wood decomposer communities. Ecology Letters 13: 675–684.

Galand, P. E., E. O. Casamayor, D. L. Kirchman, M. Potvin, & C. Lovejoy, 2009. Unique archaeal assemblages in the Arctic Ocean unveiled by massively parallel tag sequencing. ISME journal 3(7): 860–869.

Garnier, S., 2017. viridis: Default Color Maps from "matplotlib.". R package, https://cran.r-project.org/package=viridis.

Gentile, A. R., S. Barceló-Cordón, & M. Van Liedekerke, 2009. Soil Country Analyses - Belgium. https://publications.jrc.ec.europa.eu/repository/bitstream/JRC53315/publication_scr_belgium_final.pdf.

Ger, K. A., L. A. Hansson, & M. Lürling, 2014. Understanding cyanobacteria-zooplankton interactions in a more eutrophic world. Freshwater Biology 59: 1783–1798.

Ghai, R., C. M. Mizuno, A. Picazo, A. Camacho, & F. Rodriguez-Valera, 2014. Key roles for freshwater Actinobacteria revealed by deep metagenomic sequencing. Molecular Ecology 23: 6073–6090.

Gianuca, A. T., J. Engelen, K. I. Brans, F. T. T. Hanashiro, M. Vanhamel, E. M. van den Berg, C. Souffreau, & L. De Meester, 2018. Taxonomic, functional and phylogenetic metacommunity ecology of cladoceran zooplankton along urbanization gradients. Ecography 41: 183–194.

Gilbert, B., & J. R. Bennett, 2010. Partitioning variation in ecological communities: Do the numbers add up?. Journal of Applied Ecology 47: 1071–1082.

Goberna, M., J. A. Navarro-Cano, A. Valiente-Banuet, C. García, & M. Verdú, 2014. Abiotic stress tolerance and competition-related traits underlie phylogenetic clustering in soil bacterial communities. Ecology Letters 17: 1191–1201.

Gómez-Rodríguez, C., & A. Baselga, 2018. Variation among European beetle taxa in patterns

of distance decay of similarity suggests a major role of dispersal processes. Ecography 41: 1825–1834.

Green, J. L., B. J. M. Bohannan, & R. J. Whitaker, 2008. Microbial biogeography: From taxonomy to traits. Science 320: 1039–1043.

Griffiths, R. I., B. C. Thomson, P. James, T. Bell, M. Bailey, & A. S. Whiteley, 2011. The bacterial biogeography of British soils. Environmental Microbiology 13: 1642–1654.

Grimm, N. B., S. H. Faeth, N. E. Golubiewski, C. L. Redman, J. Wu, X. Bai, & J. M. Briggs, 2008. Global Change and the Ecology of Cities. Science 319: 756–760.

Grossart, H. P., C. Dziallas, F. Leunert, & K. W. Tang, 2010. Bacteria dispersal by hitchhiking on zooplankton. Proceedings of the National Academy of Sciences of the United States of America 107: 11959–11964.

Hanashiro, F. T. T., L. De Meester, M. Vanhamel, A. T. Gianuca, S. Mukherjee, L. Verbeek, E. van den Berg, & C. Souffreau, 2022. Bacterioplankton Assembly Along a Eutrophication Gradient Is Mainly Structured by Environmental Filtering, Including Indirect Effects of Phytoplankton Composition. Microbial Ecology, https://doi.org/10.1007/s00248-022-01994-x.

Hanashiro, F. T. T., S. Mukherjee, C. Souffreau, J. Engelen, K. I. Brans, P. Busschaert, & L. De Meester, 2019. Freshwater bacterioplankton metacommunity structure along urbanization gradients in Belgium. Frontiers in Microbiology 10: 743.

Hanson, C. A., J. A. Fuhrman, M. C. Horner-Devine, & J. B. H. Martiny, 2012. Beyond biogeographic patterns: Processes shaping the microbial landscape. Nature Reviews Microbiology 10: 497–506.

Hedlund, K., I. S. Regina, W. H. Van Der Putten, J. Lepš, T. Díaz, G. W. Korthals, S. Lavorel, V. K. Brown, D. Gormsen, S. R. Mortimer, C. R. Barrueco, J. Roy, P. Smilauer, M. Smilauerová, & C. Van Dijk, 2003. Plant species diversity, plant biomass and responses of the soil community on abandoned land across Europe: Idiosyncracy or above-belowground time lags. Oikos 103: 45–58.

Heino, J., M. Tolkkinen, A. M. Pirttilä, H. Aisala, & H. Mykrä, 2014. Microbial diversity and community-environment relationships in boreal streams. Journal of Biogeography 41: 2234–2244.

Hobbs, R. J., S. Arico, J. Aronson, J. S. Baron, P. Bridgewater, V. A. Cramer, P. R. Epstein, J. J. Ewel, C. A. Klink, A. E. Lugo, D. Norton, D. Ojima, D. M. Richardson, E. W. Sanderson, F.

Valladares, M. Vilà, R. Zamora, & M. Zobel, 2006. Novel ecosystems: Theoretical and management aspects of the new ecological world order. Global Ecology and Biogeography 15: 1–7.

Horner-Devine, M. C., M. A. Leibold, V. H. Smith, & B. J. M. Bohannan, 2003. Bacterial diversity patterns along a gradient of primary productivity. Ecology Letters 6: 613–622.

Hu, A., S. Li, L. Zhang, H. Wang, J. Yang, Z. Luo, A. Rashid, S. Chen, W. Huang, & C. P. Yu, 2018. Prokaryotic footprints in urban water ecosystems: A case study of urban landscape ponds in a coastal city, China. Environmental Pollution 242: 1729–1739.

Huber, P., S. Metz, F. Unrein, G. Mayora, H. Sarmento, & M. Devercelli, 2020. Environmental heterogeneity determines the ecological processes that govern bacterial metacommunity assembly in a floodplain river system. ISME Journal 1–16.

Hug, L. A., B. J. Baker, K. Anantharaman, C. T. Brown, A. J. Probst, C. J. Castelle, C. N. Butterfield, A. W. Hernsdorf, Y. Amano, K. Ise, Y. Suzuki, N. Dudek, D. A. Relman, K. M. Finstad, R. Amundson, B. C. Thomas, & J. F. Banfield, 2016. A new view of the tree of life. Nature Microbiology 1: 1–6.

Huisman, J., G. A. Codd, H. W. Paerl, B. W. Ibelings, J. M. H. Verspagen, & P. M. Visser, 2018. Cyanobacterial blooms. Nature Reviews Microbiology 16: 471–483.

Hurley, C., 2019. gclus: Clustering Graphics. R package, https://cran.r-project.org/package =gclus.

IPCC, V. Masson-Delmotte, P. Zhai, A. Pirani, S. L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M. I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J. B. R. Matthews, T. K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, & B. Zhou, 2021. Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, https://www.ipcc.ch/%0Ahttps://www.ipcc.ch/report/ar6/wg1/.

Jansson, J. K., & K. S. Hofmockel, 2020. Soil microbiomes and climate change. Nature Reviews Microbiology 18: 35-46.

Jeraldo, P., M. Sipos, N. Chia, J. M. Brulc, A. S. Dhillon, M. E. Konkel, C. L. Larson, K. E. Nelson, A. Qu, L. B. Schook, F. Yang, B. A. White, & N. Goldenfeld, 2012. Quantification of the relative roles of niche and neutral processes in structuring gastrointestinal microbiomes. Proceedings of the National Academy of Sciences of the United States of America 109: 9692–

9698.

Johnson, P. T. J., J. T. Hoverman, V. J. McKenzie, A. R. Blaustein, & K. L. D. Richgels, 2013. Urbanization and wetland communities: Applying metacommunity theory to understand the local and landscape effects. Journal of Applied Ecology 50: 34–42.

Jones, S. E., T. A. Cadkin, R. J. Newton, & K. D. McMahon, 2012. Spatial and temporal scales of aquatic bacterial beta diversity. Frontiers in Microbiology 3: 1–10.

Jones, S. E., & K. D. McMahon, 2009. Species-sorting may explain an apparent minimal effect of immigration on freshwater bacterial community dynamics. Environmental Microbiology 11: 905–913.

Ju, F., Y. Xia, F. Guo, Z. Wang, & T. Zhang, 2014. Taxonomic relatedness shapes bacterial assembly in activated sludge of globally distributed wastewater treatment plants. Environmental Microbiology 16: 2421–2432.

Jürgens, K., H. Arndt, & K. O. Rothhaupt, 1994. Zooplankton-mediated changes of bacterial community structure. Microbial Ecology 27: 27–42.

Jürgens, K., & E. Jeppesen, 2000. The impact of metazooplankton on the structure of the microbial food web in a shallow, hypertrophic lake. Journal of Plankton Research 22: 1047–1070.

Kampe, H., C. Dziallas, H. Grossart, & N. Kamjunke, 2010. Similar Bacterial Community Composition in Acidic Mining Lakes with Different pH and Lake Chemistry. Microbial Ecology 60: 618–627.

Kasprzak, P., J. Padisák, R. Koschel, L. Krienitz, & F. Gervais, 2008. Chlorophyll a concentration across a trophic gradient of lakes: An estimator of phytoplankton biomass?. Limnologica 38: 327–338.

Kaye, J. P., P. M. Groffman, N. B. Grimm, L. A. Baker, & R. V. Pouyat, 2006. A distinct urban biogeochemistry?. Trends in Ecology & Evolution 21: 192–199.

Kiersztyn, B., R. Chróst, T. Kaliński, W. Siuda, A. Bukowska, G. Kowalczyk, & K. Grabowska, 2019. Structural and functional microbial diversity along a eutrophication gradient of interconnected lakes undergoing anthropopressure. Scientific Reports 9(1): 11144.

Kirk, K. L., 1991. Inorganic Particles Alter Competition in Grazing Plankton : The Role of Selective Feeding. Ecology 72: 915–923.

Klindworth, A., E. Pruesse, T. Schweer, J. Peplies, C. Quast, M. Horn, & F. O. Glöckner, 2013. Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. Nucleic Acids Research 41(1): e1.

Knight, R., A. Vrbanac, B. C. Taylor, A. Aksenov, C. Callewaert, J. Debelius, A. Gonzalez, T. Kosciolek, L. I. McCall, D. McDonald, A. V. Melnik, J. T. Morton, J. Navas, R. A. Quinn, J. G. Sanders, A. D. Swafford, L. R. Thompson, A. Tripathi, Z. Z. Xu, J. R. Zaneveld, Q. Zhu, J. G. Caporaso, & P. C. Dorrestein, 2018. Best practices for analysing microbiomes. Nature Reviews Microbiology 16: 410–422.

Kolde, R., 2019. pheatmap: Pretty Heatmaps. R package, https://cran.r-project.org/package =pheatmap

Konopka, A., 2009. What is microbial community ecology. ISME Journal 3: 1223–1230.

Kosten, S., V. L. M. Huszar, E. Bécares, L. S. Costa, E. van Donk, L. A. Hansson, E. Jeppesen, C. Kruk, G. Lacerot, N. Mazzeo, L. De Meester, B. Moss, M. Lürling, T. Nõges, S. Romo, & M. Scheffer, 2012. Warmer climates boost cyanobacterial dominance in shallow lakes. Global Change Biology 18: 118–126.

Kozak, K. H., C. H. Graham, & J. J. Wiens, 2008. Integrating GIS-based environmental data into evolutionary biology. Trends in Ecology & Evolution 23: 141–148.

Krause, S., X. Le Roux, P. A. Niklaus, P. M. van Bodegom, J. T. Lennon T., S. Bertilsson, H. P. Grossart, L. Philippot, & P. L. E. Bodelier, 2014. Trait-based approaches for understanding microbial biodiversity and ecosystem functioning. Frontiers in Microbiology 5: 1–10.

Kremen, C., 2005. Managing ecosystem services: what do we need to know about their ecoclogy?. Ecology Letters 8: 468–479.

Lange, V., I. Böhme, J. Hofmann, K. Lang, J. Sauter, B. Schöne, P. Paul, V. Albrecht, J. M. Andreas, D. M. Baier, J. Nething, U. Ehninger, C. Schwarzelt, J. Pingel, G. Ehninger, & A. H. Schmidt, 2014. Cost-efficient high-throughput HLA typing by MiSeq amplicon sequencing. BMC Genomics 15: 63.

Langenheder, S., M. Berga, Ö. Östman, & A. J. Székely, 2012. Temporal variation of B-diversity and assembly mechanisms in a bacterial metacommunity. ISME Journal 6: 1107–1114.

Langenheder, S., & K. Jürgens, 2001. Regulation of bacterial biomass and community structure by metazoan and protozoan predation. Limnology and Oceanography 46: 121–134.

Langenheder, S., & E. S. Lindström, 2019. Factors influencing aquatic and terrestrial bacterial community assembly. Environmental Microbiology Reports 11: 306–315.

Langenheder, S., & A. J. Székely, 2011. Species sorting and neutral processes are both important during the initial assembly of bacterial communities. ISME Journal 5: 1086–1094.

Langille, M. G. I., J. Zaneveld, J. G. Caporaso, D. McDonald, D. Knights, J. A. Reyes, J. C. Clemente, D. E. Burkepile, R. L. Vega Thurber, R. Knight, R. G. Beiko, & C. Huttenhower, 2013. Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. Nature Biotechnology 31: 814–821.

Lauber, C. L., K. S. Ramirez, Z. Aanderud, J. Lennon, & N. Fierer, 2013. Temporal variability in soil microbial communities across land-use types. ISME Journal 7: 1641–1650.

Lawton, J. H., 1999. Size matters. Oikos 85: 19–21.

Legendre, P., & M. J. Andersson, 1999. Distance-based redundancy analysis: Testing multispecies responses in multifactorial ecological experiments. Ecological Monographs 69: 1–24.

Legendre, P., & E. D. Gallagher, 2001. Ecologically meaningful transformations for ordination of species data. Oecologia 129: 271–280.

Leibold, M. A., M. Holyoak, N. Mouquet, P. Amarasekare, J. M. Chase, M. F. Hoopes, R. D. Holt, J. B. Shurin, R. Law, D. Tilman, M. Loreau, & A. Gonzalez, 2004. The metacommunity concept: a framework for multi-scale community ecology. Ecology Letters 7: 601–613.

Leibold, M. A., & G. M. Mikkelson, 2002. Coherence, species turnover, and boundary clumping elements of meta-community structure. Oikos 97: 237–250.

Lewis, S. L., & M. A. Maslin, 2015. Defining the Anthropocene. Nature 519: 171–180.

Liao, J., X. Cao, J. Wang, L. Zhao, J. Sun, D. Jiang, & Y. Huang, 2017. Similar community assembly mechanisms underlie similar biogeography of rare and abundant bacteria in lakes on Yungui Plateau, China. Limnology and Oceanography 62: 723–735.

Lima-Mendez, G., K. Faust, N. Henry, J. Decelle, S. Colin, F. Carcillo, S. Chaffron, J. C. Ignacio-espinosa, S. Roux, F. Vincent, & L. Bittner, 2015. Determinants of community structure in the global plankton interactome. Science 348: 1–10.

Lima, M. S., D. da Motta Marques, N. H. They, K. D. McMahon, L. R. Rodrigues, L. de S. Cardoso, & L. O. Crossetti, 2016. Contrasting factors drive within-lake bacterial community

composition and functional traits in a large shallow subtropical lake. Hydrobiologia 778: 105–120.

Lindström, E. S., & S. Langenheder, 2012. Local and regional factors influencing bacterial community assembly. Environmental Microbiology Reports 4: 1–9.

Liu, L., J. Yang, Z. Yu, & D. M. Wilkinson, 2015. The biogeography of abundant and rare bacterioplankton in the lakes and reservoirs of China. ISME Journal 9: 2068–2077.

Logares, R., E. S. Lindström, S. Langenheder, J. B. Logue, H. Paterson, J. Laybourn-Parry, K. Rengefors, L. Tranvik, & S. Bertilsson, 2013. Biogeography of bacterial communities exposed to progressive long-term environmental change. ISME Journal 7: 937–948.

Logue, J. B., & E. S. Lindström, 2008. Biogeography of Bacterioplankton in Inland Waters. Freshwater Reviews 1: 99–114.

Logue, J. B., & E. S. Lindström, 2010. Species sorting affects bacterioplankton community composition as determined by 16S rDNA and 16S rRNA fingerprints. ISME Journal 4: 729–738.

Logue, J. B., N. Mouquet, H. Peter, & H. Hillebrand, 2011. Empirical approaches to metacommunities: A review and comparison with theory. Trends in Ecology & Evolution 26: 482–491.

Louca, S., S. M. S. Jacques, A. P. F. Pires, J. S. Leal, A. L. González, M. Doebeli, & V. F. Farjalla, 2017. Functional structure of the bromeliad tank microbiome is strongly shaped by local geochemical conditions. Environmental Microbiology 19: 3132–3151.

Louca, S., L. W. Parfrey, & M. Doebeli, 2016. Decoupling function and taxonomy in the global ocean microbiome. Science 353: 1272–1277.

Lu, S., S. Gischkat, M. Reiche, D. M. Akob, K. B. Hallberg, & K. Küsel, 2010. Ecophysiology of Fe-Cycling Bacteria in Acidic Sediments. Applied and Environmental Microbiology 76: 8174–8183.

Lundin, D., I. Severin, J. B. Logue, Ö. Östman, A. F. Andersson, & E. S. Lindström, 2012. Which sequencing depth is sufficient to describe patterns in bacterial α- and β-diversity?. Environmental Microbiology Reports 4: 367–372.

Luo, C., D. Tsementzi, N. Kyrpides, T. Read, & K. T. Konstantinidis, 2012. Direct comparisons of Illumina vs. Roche 454 sequencing technologies on the same microbial community DNA

sample. PLoS ONE 7(2): e30087.

Lynch, M. D. J., & J. D. Neufeld, 2015. Ecology and exploration of the rare biosphere. Nature Reviews Microbiology 13: 217–229.

Martiny, A. C., A. P. K. Tai, D. Veneziano, F. Primeau, & S. W. Chisholm, 2009. Taxonomic resolution, ecotypes and the biogeography of *Prochlorococcus*. Environmental Microbiology 11: 823–832.

Martiny, A. C., K. Treseder, & G. Pusch, 2013. Phylogenetic conservatism of functional traits in microorganisms. ISME Journal 7: 830–838,.

Martiny, J. B. H., B. J. M. Bohannan, J. H. Brown, R. K. Colwell, J. A. Fuhrman, J. L. Green, M. C. Horner-Devine, M. Kane, J. A. Krumins, C. R. Kuske, P. J. Morin, S. Naeem, L. Øvreås, A. L. Reysenbach, V. H. Smith, & J. T. Staley, 2006. Microbial biogeography: Putting microorganisms on the map. Nature Reviews Microbiology 4: 102–112.

Martiny, J. B. H., S. E. Jones, J. T. Lennon, & A. C. Martiny, 2015. Microbiomes in light of traits: A phylogenetic perspective. Science 350(6261): aac9323.

McDonnell, M. J., & A. K. Hahs, 2015. Adaptation and Adaptedness of Organisms to Urban Environments. Annual Review of Ecology, Evolution, and Systematics 46: 261–280.

McKinney, M. L., 2006. Urbanization as a major cause of biotic homogenization. Biological Conservation 127: 247–260.

McKinney, M. L., 2008. Effects of urbanization on species richness: A review of plants and animals. Urban Ecosystems 11: 161–176.

McMurdie, P. J., & S. Holmes, 2013. Phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. PLoS ONE 8(4): e61217.

Meersmans, J., B. Van Wesemael, E. Goidts, M. Van Molle, S. De Baets, & F. De Ridder, 2011. Spatial analysis of soil organic carbon evolution in Belgian croplands and grasslands, 1960-2006. Global Change Biology 17: 466–479.

Menezes, S., D. J. Baird, & A. M. V. M. Soares, 2010. Beyond taxonomy: a review of macroinvertebrate trait-based community descriptors as tools for freshwater biomonitoring. Journal of Applied Ecology 47: 711–719.

Merckx, T., C. Souffreau, A. Kaiser, L. F. Baardsen, T. Backeljau, D. Bonte, K. I. Brans, M. Cours, M. Dahirel, N. Debortoli, K. De Wolf, J. M. T. Engelen, D. Fontaneto, A. T. Gianuca,

L. Govaert, F. Hendrickx, J. Higuti, L. Lens, K. Martens, H. Matheve, E. Matthysen, E. Piano, R. Sablon, I. Schön, K. Van Doninck, L. De Meester, & H. Van Dyck, 2018. Body-size shifts in aquatic and terrestrial urban communities. Nature 558: 113–116.

Møller, A. P., 2010. The fitness benefit of association with humans: Elevated success of birds breeding indoors. Behavioral Ecology 21: 913–918.

Monard, C., S. Gantner, S. Bertilsson, S. Hallin, & J. Stenlid, 2016. Habitat generalists and specialists in microbial communities across a terrestrial-freshwater gradient. Scientific Reports 6: 1–10.

Monchamp, M. E., P. Spaak, I. Domaizon, N. Dubois, D. Bouffard, & F. Pomati, 2018. Homogenization of lake cyanobacterial communities over a century of climate change and eutrophication. Nature Ecology & Evolution 2: 317–324.

Mony, C., P. Vandenkoornhuyse, B. J. M. Bohannan, K. Peay, & M. A. Leibold, 2020. A Landscape of Opportunities for Microbial Ecology Research. Frontiers in Microbiology 11: 561427.

Moritz, C., J. L. Patton, C. J. Conroy, J. L. Parra, G. C. White, & S. R. Beissinger, 2008. Impact of a century of climate change on small-mammal communities in Yosemite National Park, USA. Science 322: 261–264.

Morse, C. C., A. D. Huryn, & C. Cronan, 2003. Impervious surface area as a predictor of the effects of urbanization on stream insect communities in Maine, U.S.A. Environmental Monitoring and Assessment 89: 95–127.

Mukherjee, S., H. Juottonen, P. Siivonen, C. Lloret Quesada, P. Tuomi, P. Pulkkinen, & K. Yrjälä, 2014. Spatial patterns of microbial diversity and activity in an aged creosote-contaminated site. ISME Journal 8: 2131–2142.

Murphy, J., & J. P. Riley, 1962. A modified single solution method for the determination of phosphate in natural waters. Analytica Chimica Acta 27: 31–36.

Muscarella, M. E., S. E. Jones, & J. T. Lennon, 2016. Species sorting along a subsidy gradient alters bacterial community stability. Ecology 97: 2034–2043.

Mykrä, H., M. Tolkkinen, & J. Heino, 2017. Environmental degradation results in contrasting changes in the assembly processes of stream bacterial and fungal communities. Oikos 126: 1291–1298.

Nekola, J. C., & P. S. White, 1999. The distance decay of similarity in biogeography and ecology. Journal of Biogeography 26: 867–878.

Neuenschwander, S. M., J. Pernthaler, T. Posch, & M. M. Salcher, 2015. Seasonal growth potential of rare lake water bacteria suggest their disproportional contribution to carbon fluxes. Environmental Microbiology 17: 781–795.

Neuwirth, E., 2014. RColorBrewer: ColorBrewer Palettes. R package, https://cran.r-project.org /package=RColorBrewer.

Newton, R. J., S. E. Jones, A. Eiler, K. D. McMahon, & S. Bertilsson, 2011. A Guide to the Natural History of Freshwater Lake Bacteria. Microbiology and Molecular Biology Reviews 75: 14–49.

Newton, R. J., & S. L. McLellan, 2015. A unique assemblage of cosmopolitan freshwater bacteria and higher community diversity differentiate an urbanized estuary from oligotrophic Lake Michigan. Frontiers in Microbiology 6: 1–13.

O' Reilly, C. M., R. J. Rowley, P. Schneider, J. D. Lenters, P. B. Mcintyre, B. M. Kraemer, G. A. Weyhenmeyer, D. Straile, B. Dong, R. Adrian, M. G. Allan, O. Anneville, L. Arvola, J. Austin, J. L. Bailey, J. S. Baron, J. D. Brookes, E. de Eyto, M. T. Dokulil, D. P. Hamilton, K. Havens, A. L. Hetherington, S. N. Higgins, S. Hook, L. R. Izmest'eva, K. D. Joehnk, K. Kangur, P. Kasprzak, M. Kumagai, E. Kuusisto, G. Leshkevich, D. M. Livingstone, S. MacIntyre, L. May, J. M. Melack, D. C. Mueller-Navarra, M. Naumenko, P. Noges, T. Noges, R. P. North, P.-D. Plisnier, A. Rigosi, A. Rimmer, M. Rogora, L. G. Rudstam, J. A. Rusak, N. Salmaso, N. R. Samal, D. E. Schindler, S. G. Schladow, M. Schmid, S. R. Schmidt, E. Silow, M. E. Soylu, K. Teubner, P. Verburg, A. Voutilainen, A. Watkinson, C. E. Williamson, & G. Zhang, 2015. Rapid and highly variable warming of lake surface waters around the globe. Geophysical Research Letters 42(24): 10773-10781.

O'Malley, M. A, 2007. The nineteenth century roots of "everything is everywhere." Nature Reviews Microbiology 5: 647–651.

Oksanen, J., 2015. Vegan: Community Ecology Package. version 2.3–0. Rpackage, http:// cc.oulu.fi/jarioksa/softhelp/vegan.html.

Östman, Ö., S. Drakare, E. S. Kritzberg, S. Langenheder, J. B. Logue, & E. S. Lindström, 2010. Regional invariance among microbial communities. Ecology Letters 13: 118–127.

Ovaskainen, O., G. Tikhonov, A. Norberg, F. Guillaume Blanchet, L. Duan, D. Dunson, T.

Roslin, & N. Abrego, 2017. How to make more out of community data? A conceptual framework and its implementation as models and software. Ecology Letters 20: 561–576.

Özen, A., M. Šorf, C. Trochine, L. Liboriussen, M. Beklioglu, M. Søndergaard, T. L. Lauridsen, L. S. Johansson, & E. Jeppesen, 2013. Long-term effects of warming and nutrients on microbes and other plankton in mesocosms. Freshwater Biology 58: 483–493.

Paerl, H. W., & J. Huisman, 2008. Blooms like it hot. Science 320: 57–58.

Paradis, E., J. Claude, & K. Strimmer, 2004. APE: Analyses of phylogenetics and evolution in R language. Bioinformatics 20: 289–290.

Parmesan, C., & G. Yohe, 2003. A globally coherent fingerprint of climate change impacts across natural systems. Nature 421: 37–42.

Parveen, B., J. P. Reveilliez, I. Mary, V. Ravet, G. Bronner, J. F. Mangot, I. Domaizon, & D. Debroas, 2011. Diversity and dynamics of free-living and particle-associated Betaproteobacteria and Actinobacteria in relation to phytoplankton and zooplankton communities. FEMS Microbiology Ecology 77: 461–476.

Pedrós-Alió, C., 2012. The Rare Bacterial Biosphere. Annual Review of Marine Science 4: 449–466.

Peres-Neto, P. R., P. Legendre, S. Dray, & D. Borcard, 2006. Variation partitioning of species data matrices: Estimation and comparison of fractions. Ecology 87: 2614–2625.

Pernthaler, J., 2005. Predation on prokaryotes in the water column and its ecological implications. Nature Reviews Microbiology 3: 537–546.

Piano, E., K. De Wolf, F. Bona, D. Bonte, D. E. Bowler, M. Isaia, L. Lens, T. Merckx, D. Mertens, M. van Kerckvoorde, L. De Meester, & F. Hendrickx, 2017. Urbanization drives community shifts towards thermophilic and dispersive species at local and landscape scales. Global Change Biology 23: 2554–2564.

Pollock, J., L. Glendinning, T. Wisedchanwet, & M. Watson, 2018. The madness of microbiome: Attempting to find consensus “best practice” for 16S microbiome studies. Applied and Environmental Microbiology 84(7): e02627-17.

Powell, J. R., S. Karunaratne, C. D. Campbell, H. Yao, L. Robinson, & B. K. Singh, 2015. Deterministic processes vary during community assembly for ecologically dissimilar taxa. Nature Communications 6: 8444.

Preston-Mafham, J., L. Boddy, & P. F. Randerson, 2002. Analysis of microbial community functional diversity using sole-carbon-source utilisation profiles - A critique. FEMS Microbiology Ecology 42: 1–14.

Prosser, J. I., B. J. M. Bohannan, T. P. Curtis, R. J. Ellis, M. K. Firestone, R. P. Freckleton, J. L. Green, L. E. Green, K. Killham, J. J. Lennon, a M. Osborn, M. Solan, C. J. van der Gast, & J. P. W. Young, 2007. The role of ecological theory in microbial ecology. Nature Reviews Microbiology 5: 384–392.

Qi, W., G. Nong, J. F. Preston, F. Ben-Ami, & D. Ebert, 2009. Comparative metagenomics of Daphnia symbionts. BMC Genomics 10: 172.

R Core Team, 2018. R: A Language and Environment for Statistical Computing. Vienna, Austria, https://www.r-project.org/.

Ram, A. S. P., S. Palesse, J. Colombet, A. Thouvenot, & T. Sime-Ngando, 2014. The relative importance of viral lysis and nanoflagellate grazing for prokaryote mortality in temperate lakes. Freshwater Biology 59: 300–311.

Ramirez, K. S., J. W. Leff, A. Barberán, S. T. Bates, J. Betley, T. W. Crowther, E. F. Kelly, E. E. Oldfield, E. Ashley Shaw, C. Steenbock, M. A. Bradford, D. H. Wall, & N. Fierer, 2014. Biogeographic patterns in below-ground diversity in New York City's Central Park are similar to those observed globally. Proceedings of the Royal Society B: Biological Sciences 281: 20141988.

Reese, A. T., A. Savage, E. Youngsteadt, K. L. Mcguire, A. Koling, O. Watkins, S. D. Frank, & R. R. Dunn, 2016. Urban stress is associated with variation in microbial species composition - but not richness - in Manhattan. ISME Journal 10: 751–760.

Ren, L., D. He, Z. Chen, E. Jeppesen, T. L. Lauridsen, M. Søndergaard, Z. Liu, & Q. L. Wu, 2017. Warming and nutrient enrichment in combination increase stochasticity and beta diversity of bacterioplankton assemblages across freshwater mesocosms. ISME Journal 11: 613–625.

Ren, L., D. He, J. Zeng, & Q. L. Wu, 2013. Bacterioplankton communities turn unstable and become small under increased temperature and nutrient-enriched conditions. FEMS Microbiology Ecology 84: 614–624.

Ren, L., E. Jeppesen, D. He, J. Wang, L. Liboriussen, P. Xing, & Q. L. Wu, 2015. pH influences the importance of niche-related and neutral processes in lacustrine bacterioplankton assembly.

Applied and Environmental Microbiology 81: 3104–3114.

Rigosi, A., C. C. Carey, B. W. Ibelings, & J. D. Brookes, 2014. The interaction between climate warming and eutrophication to promote cyanobacteria is dependent on trophic state and varies among taxa. Limnology and Oceanography 59: 99–144.

Roguet, A., G. S. Laigle, C. Therial, A. Bressy, F. Soulignac, A. Catherine, G. Lacroix, L. Jardillier, C. Bonhomme, T. Z. Lerch, & F. S. Lucas, 2015. Neutral community model explains the bacterial community assembly in freshwater lakes. FEMS Microbiology Ecology 91: 1–11.

Rosseel, Y., 2012. {lavaan}: An {R} Package for Structural Equation Modeling. Journal of Statistical Software 48: 1–36.

Rothschild, L. J., & R. L. Mancinelli, 2001. Life in extreme environments. Nature 409: 1092–1101.

Ruiz-González, C., J. P. Niño-García, & P. A. del Giorgio, 2015a. Terrestrial origin of bacterial communities in complex boreal freshwater networks. Ecology Letters 18: 1198–1206.

Ruiz-González, C., J. P. Niño-García, J. F. Lapierre, & P. A. del Giorgio, 2015b. The quality of organic matter shapes the functional biogeography of bacterioplankton across boreal freshwater ecosystems. Global Ecology and Biogeography 24: 1487–1498.

Rummens, K., L. De Meester, & C. Souffreau, 2018. Inoculation history affects community composition in experimental freshwater bacterioplankton communities. Environmental Microbiology 20: 1120–1133.

Rutgers, M., M. Wouterse, S. M. Drost, A. M. Breure, C. Mulder, D. Stone, R. E. Creamer, A. Winding, & J. Bloem, 2016. Monitoring soil bacteria with community-level physiological profiles using Biolog^{TM} ECO-plates in the Netherlands and Europe. Applied Soil Ecology 97: 23–35.

Saito, M., & F. Koike, 2013. Distribution of Wild Mammal Assemblages along an Urban-Rural-Forest Landscape Gradient in Warm-Temperate East Asia. PLoS ONE 8: e65464.

Scheffer, M., S. Carpenter, J. A. Foley, C. Folke, & B. Walker, 2001. Catastrophic shifts in ecosystems. Nature 413: 591–596.

Scheffer, M., & E. H. Van Nes, 2007. Shallow lakes theory revisited: Various alternative regimes driven by climate, nutrients, depth and lake size. Hydrobiologia 584: 455–466.

Schindler, D. W., 2012. The dilemma of controlling cultural eutrophication of lakes.

Proceedings of the Royal Society B: Biological Sciences 279: 4322–4333.

Schloerke, B., J. Crowley, D. Cook, F. Briatte, M. Marbach, E. Thoen, A. Elberg, & J. Larmarange, 2018. GGally: Extension to "ggplot2". R package, https://cran.r-project.org/package=GGally.

Schloss, P. D., D. Gevers, & S. L. Westcott, 2011. Reducing the effects of PCR amplification and sequencing artifacts on 16S rRNA-based studies. PloS ONE 6: e27310.

Schloss, P. D., S. L. Westcott, T. Ryabin, J. R. Hall, M. Hartmann, E. B. Hollister, R. A. Lesniewski, B. B. Oakley, D. H. Parks, C. J. Robinson, J. W. Sahl, B. Stres, G. G. Thallinger, D. J. Van Horn, & C. F. Weber, 2009. Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. Applied and Environmental Microbiology 75: 7537–7541.

Shade, A., R. R. Dunn, S. A. Blowes, P. Keil, B. J. M. Bohannan, M. Herrmann, K. Küsel, J. T. Lennon, N. J. Sanders, D. Storch, & J. Chase, 2018. Macroecology to Unite All Life, Large and Small. Trends in Ecology & Evolution 33: 731–744.

Shade, A., J. S. Read, N. D. Youngblut, N. Fierer, R. Knight, T. K. Kratz, N. R. Lottig, E. E. Roden, E. H. Stanley, J. Stombaugh, R. J. Whitaker, C. H. Wu, & K. D. McMahon, 2012. Lake microbial communities are resilient after a whole-ecosystem disturbance. ISME Journal 6: 2153–2167.

Shi, Y., P. Grogan, H. Sun, J. Xiong, Y. Yang, J. Zhou, & H. Chu, 2015. Multi-scale variability analysis reveals the importance of spatial distance in shaping Arctic soil microbial functional communities. Soil Biology and Biochemistry 86: 126–134.

Shurin, J. B., J. L. Clasen, H. S. Greig, P. Kratina, & P. L. Thompson, 2012. Warming shifts top-down and bottom-up control of pond food web structure and function. Philosophical Transactions of the Royal Society B: Biological Sciences 367: 3008–3017.

Šimek, K., J. Nedoma, P. Znachor, V. Kasalický, J. Jezbera, K. Horňák, & J. Sed'a, 2014. A finely tuned symphony of factors modulates the microbial food web of a freshwater reservoir in spring. Limnology and Oceanography 59: 1477–1492.

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

Sinclair, L., O. A. Osman, S. Bertilsson, & A. Eiler, 2015. Microbial community composition and diversity via 16S rRNA gene amplicons: Evaluating the Illumina platform. PLoS ONE 10: 1–18.

Siqueira, T., L. M. Bini, F. O. Roque, S. R. Marques Couceiro, S. Trivinho-Strixino, & K. Cottenie, 2012. Common and rare species respond to similar niche processes in macroinvertebrate metacommunities. Ecography 35: 183–192.

Sison-Mangus, M. P., A. A. Mushegian, & D. Ebert, 2015. Water fleas require microbiota for survival, growth and reproduction. ISME Journal 9: 59–67.

Smith, E. M., & Y. T. Prairie, 2004. Bacterial metabolism and growth efficiency in lakes: The importance of phosphorus availability. Limnology and Oceanography 49: 137–147.

Smith, V. H., 2007. Microbial diversity-productivity relationships in aquatic ecosystems. FEMS Microbiology Ecology 62: 181–186.

Soininen, J., 2014. A quantitative analysis of species sorting across organisms and ecosystems. Ecology 95: 3284–3292.

Soininen, J., J. J. Korhonen, J. Karhu, & A. Vetterli, 2011. Disentangling the spatial patterns in community composition of prokaryotic and eukaryotic lake plankton. Limnology and Oceanography 56: 508–520.

Souffreau, C., P. Busschaert, C. Denis, J. Van Wichelen, B. Lievens, W. Vyverman, & L. De Meester, 2018. A comparative hierarchical analysis of bacterioplankton and biofilm metacommunity structure in an interconnected pond system. Environmental Microbiology 20: 1271–1282.

Souffreau, C., B. Pecceu, C. Denis, K. Rummens, & L. De Meester, 2014. An experimental analysis of species sorting and mass effects in freshwater bacterioplankton. Freshwater Biology 59: 2081–2095.

Souffreau, C., K. Van der Gucht, I. van Gremberghe, S. Kosten, G. Lacerot, L. M. Lobão, V. L. de Moraes Huszar, F. Roland, E. Jeppesen, W. Vyverman, & L. De Meester, 2015. Environmental rather than spatial factors structure bacterioplankton communities in shallow lakes along a >6000km latitudinal gradient in South America. Environmental Microbiology 17: 2336–2351.

Stegen, J. C., X. Lin, J. K. Fredrickson, X. Chen, D. W. Kennedy, C. J. Murray, M. L. Rockhold, & A. Konopka, 2013. Quantifying community assembly processes and identifying features that

impose them. ISME Journal 7: 2069–2079.

Stegen, J. C., X. Lin, A. E. Konopka, & J. K. Fredrickson, 2012. Stochastic and deterministic assembly processes in subsurface microbial communities. ISME Journal 6: 1653–1664.

Stocker, T. F., D. Qin, G. K. Plattner, M. M. B. Tignor, S. K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, & P. M. Midgley, 2013. Climate change 2013 the physical science basis: Working Group I contribution to the fifth assessment report of the intergovernmental panel on climate change. Climate Change 2013 the Physical Science Basis: Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press.

Su, X., A. D. Steinman, X. Tang, Q. Xue, Y. Zhao, & L. Xie, 2017. Response of bacterial communities to cyanobacterial harmful algal blooms in Lake Taihu, China. Harmful Algae 68: 168–177.

Székely, A. J., & S. Langenheder, 2014. The importance of species sorting differs between habitat generalists and specialists in bacterial communities. FEMS Microbiology Ecology 87: 102–112.

Tamburini, C., M. Boutrif, M. Garel, R. R. Colwell, & J. W. Deming, 2013. Prokaryotic responses to hydrostatic pressure in the ocean - a review. Environmental Microbiology 15: 1262–1274.

Tang, Y., M. Horikoshi, & W. Li, 2016. ggfortify: Unified interface to visualize statistical results of popular r packages. R Journal 8: 478–489.

Tank, S. E., L. F. W. Lesack, & D. J. Mcqueen, 2009. Elevated pH regulates bacterial carbon cycling in lakes with high photosynthetic activity. Ecology 90: 1910–1922.

Teittinen, A., M. Taka, O. Ruth, & J. Soininen, 2015. Variation in stream diatom communities in relation to water quality and catchment variables in a boreal, urbanized region. Science of the Total Environment 530–531: 279–289.

Tilman, D., M. Clark, D. R. Williams, K. Kimmel, S. Polasky, & C. Packer, 2017. Future threats to biodiversity and pathways to their prevention. Nature 546: 73–81.

Tiquia, S. M., 2010. Metabolic diversity of the heterotrophic microorganisms and potential link to pollution of the Rouge River. Environmental Pollution 158: 1435–1443.

Tripathi, B. M., J. C. Stegen, M. Kim, K. Dong, J. M. Adams, & Y. K. Lee, 2018. Soil pH

mediates the balance between stochastic and deterministic assembly of bacteria. ISME Journal 12: 1072–1083.

United States Environmental Protection Agency (USEPA), 1983. Methods of chemical analysis for water and wastes. Cincinatti, OH, USA.

Vaas, L. A. I., J. Sikorski, B. Hofner, A. Fiebig, N. Buddruhs, H. P. Klenk, & M. Göker, 2013. Opm: An R package for analysing OmniLog® phenotype microarray data. Bioinformatics 29: 1823–1824.

Valverde, A., T. P. Makhalanyane, & D. A. Cowan, 2014. Contrasting assembly processes in a bacterial metacommunity along a desiccation gradient. Frontiers in Microbiology 5: 1–8.

Valverde, A., T. P. Makhalanyane, M. Seely, & D. A. Cowan, 2015. Cyanobacteria drive community composition and functionality in rock-soil interface communities. Molecular Ecology 24: 812–821.

Van Der Gucht, K., K. Cottenie, K. Muylaert, N. Vloemans, S. Cousin, S. Declerck, E. Jeppesen, J. M. Conde-Porcuna, K. Schwenk, G. Zwart, H. Degans, W. Vyverman, & L. De Meester, 2007. The power of species sorting: Local factors drive bacterial community composition over a wide range of spatial scales. Proceedings of the National Academy of Sciences of the United States of America 104: 20404–20409.

Van Der Heijden, M. G. A., R. D. Bardgett, & N. M. Van Straalen, 2008. The unseen majority: Soil microbes as drivers of plant diversity and productivity in terrestrial ecosystems. Ecology Letters 11: 296–310.

Vavrek, M. J., 2011. fossil: Palaeoecological and palaeogeographical analysis tools. Palaeontologia Electronica 14: 1T.

Venail, P. A., & M. J. Vives, 2013. Phylogenetic distance and species richness interactively affect the productivity of bacterial communities. Ecology 94: 2529–2536.

Verbeek, L., M. Vanhamel, E. van den Berg, F. T. T. Hanashiro, A. T. Gianuca, M. Striebel, P. Lemmens, S. A. J. Declerck, H. Hillebrand, & L. De Meester, 2018. Compositional and functional consequences of environmental change in Belgian farmland ponds. Freshwater Biology 63: 581–596.

Verreydt, D., L. De Meester, E. Decaestecker, M. J. Villena, K. Van Der Gucht, P. Vannormelingen, W. Vyverman, & S. A. J. Declerck, 2012. Dispersal-mediated trophic interactions can generate apparent patterns of dispersal limitation in aquatic metacommunities.

Ecology Letters 15: 218–226.

Vilar, A. G., H. Van Dam, E. E. Van Loon, J. A. Vonk, H. G. Van Der Geest, & W. Admiraal, 2014. Eutrophication decreases distance decay of similarity in diatom communities. Freshwater Biology 59: 1522–1531.

Vonlanthen, P., D. Bittner, A. G. Hudson, K. A. Young, R. Müller, B. Lundsgaard-Hansen, D. Roy, S. Di Piazza, C. R. Largiader, & O. Seehausen, 2012. Eutrophication causes speciation reversal in whitefish adaptive radiations. Nature 482: 357–362.

Vörösmarty, C. J., P. B. McIntyre, M. O. Gessner, D. Dudgeon, A. Prusevich, P. Green, S. Glidden, S. E. Bunn, C. A. Sullivan, C. Reidy Liermann, & P. M. Davies, 2010. Global threats to human water security and river biodiversity. Nature 467: 555–561.

Waide, R. B., M. R. Willig, C. F. Steiner, G. Mittelbach, L. Gough, S. I. Dodson, G. P. Juday, & R. Parmenter, 1999. The relationship between productivity and species richness. Annual Review of Ecology and Systematics 30: 257–300.

Walker, S. C., & H. Cyr, 2007. Testing the standard neutral model of biodiversity in lake communities. Oikos 116: 143–155.

Walther, G. R., 2010. Community and ecosystem responses to recent climate change. Philosophical Transactions of the Royal Society B: Biological Sciences 365: 2019–2024.

Wang, S. Y., E. B. Sudduth, M. D. Wallenstein, J. P. Wright, & E. S. Bernhardt, 2011. Watershed urbanization alters the composition and function of stream bacterial communities. PLoS ONE 6: e22972.

Wang, S., D. Zhao, J. Zeng, H. Xu, R. Huang, C. Jiao, & L. Guo, 2019. Variations of bacterial community during the decomposition of Microcystis under different temperatures and biomass. BMC Microbiology BMC Microbiology 19: 1–10.

Waters, C. N., J. Zalasiewicz, C. Summerhayes, A. D. Barnosky, C. Poirier, A. Gałuszka, A. Cearreta, M. Edgeworth, E. C. Ellis, M. Ellis, C. Jeandel, R. Leinfelder, J. R. McNeill, D. D. B. Richter, W. Steffen, J. Syvitski, D. Vidas, M. Wagreich, M. Williams, A. Zhisheng, J. Grinevald, E. Odada, N. Oreskes, & A. P. Wolfe, 2016. The Anthropocene is functionally and stratigraphically distinct from the Holocene. Science 351: 6269.

Wickham, H., 2009. ggplot2 Elegant Graphics for Data Analysis. R package, http://had.co.nz /ggplot2/book.

Woodhouse, J. N., A. S. Kinsela, R. N. Collins, L. C. Bowling, G. L. Honeyman, J. K. Holliday, & B. A. Neilan, 2016. Microbial communities reflect temporal changes in cyanobacterial composition in a shallow ephemeral freshwater lake. ISME Journal 10: 1337–1351.

Wu, X., W. Xi, W. Ye, & H. Yang, 2007. Bacterial community composition of a shallow hypertrophic freshwater lake in China, revealed by 16S rRNA gene sequences. FEMS Microbiology Ecology 61: 85–96.

Xu, H. J., S. Li, J. Q. Su, S. Nie, V. Gibson, H. Li, & Y. G. Zhu, 2014. Does urbanization shape bacterial community composition in urban park soils? A case study in 16 representative Chinese cities based on the pyrosequencing method. FEMS Microbiology Ecology 87: 182–192.

Xu, H., D. Zhao, J. Zeng, C. Jiao, Z. Yu, & Q. L. Wu, 2020. Distinct successional patterns and processes of free-living and particle-attached bacterial communities throughout a phytoplankton bloom. Freshwater Biology 65: 1363–1375.

Yamamuro, M., T. Komuro, H. Kamiya, T. Kato, H. Hasegawa, & Y. Kameda, 2019. Neonicotinoids disrupt aquatic food webs and decrease fishery yields. Science 366: 620–623.

Yan, B., J. Li, N. Xiao, Y. Qi, G. Fu, G. Liu, & M. Qiao, 2016. Urban-development-induced Changes in the Diversity and Composition of the Soil Bacterial Community in Beijing. Scientific Reports 6: 38811.

Yang, Y., 2021. Emerging Patterns of Microbial Functional Traits. Trends in Microbiology 29: 874–882.

Yeh, Y. C., P. R. Peres-Neto, S. W. Huang, Y. C. Lai, C. Y. Tu, F. K. Shiah, G. C. Gong, & C. H. Hsieh, 2015. Determinism of bacterial metacommunity dynamics in the southern East China Sea varies depending on hydrography. Ecography 38: 198–212.

Zeng, J., C. Jiao, D. Zhao, H. Xu, R. Huang, X. Cao, Z. Yu, & Q. L. Wu, 2019. Patterns and assembly processes of planktonic and sedimentary bacterial community differ along a trophic gradient in freshwater lakes. Ecological Indicators 106: 105491.

Zhang, X., E. R. Johnston, W. Liu, L. Li, & X. Han, 2016. Environmental changes affect the assembly of soil bacterial community primarily by mediating stochastic processes. Global Change Biology 22: 198–207.

Zhou, J., Y. Deng, P. Zhang, K. Xue, Y. Liang, J. D. Van Nostrand, Y. Yang, Z. He, L. Wu, D. A. Stahl, T. C. Hazen, J. M. Tiedje, & A. P. Arkin, 2014. Stochasticity, succession, and environmental perturbations in a fluidic ecosystem. Proceedings of the National Academy of

Sciences of the United States of America 111(9): E836-E845.

Zhou, J., W. Liu, Y. Deng, Y. H. Jiang, K. Xue, Z. He, J. D. Van Nostrand, L. Wu, Y. Yang, & A. Wang, 2013. Stochastic assembly leads to alternative communities with distinct functions in a bioreactor microbial community. mBio 4(2): e00584-12.

Zhou, J., & D. Ning, 2017. Stochastic Community Assembly: Does It Matter in Microbial Ecology?. Microbiology and Molecular Biology Reviews 81(4): e00002-17.

Zimmerman, A. E., C. Howard-Varona, D. M. Needham, S. G. John, A. Z. Worden, M. B. Sullivan, J. R. Waldbauer, & M. L. Coleman, 2020. Metabolic and biogeochemical consequences of viral infection in aquatic ecosystems. Nature Reviews Microbiology 18(1): 21-34.

Zöllner, E., B. Santer, M. Boersma, H. G. Hoppe, & K. Jürgens, 2003. Cascading predation effects of *Daphnia* and copepods on microbial food web components. Freshwater Biology 48: 2174–2193.

Zwirglmaier, K., K. Keiz, M. Engel, J. Geist, & U. Raeder, 2015. Seasonal and spatial patterns of microbial diversity along a trophic gradient in the interconnected lakes of the Osterseen Lake District, Bavaria. Frontiers in Microbiology 6: 1168.