Doctoral theses at NTNU, 2023:403

Madeleine S. Gundersen

The effect of disturbances on bacterial community characteristics and assembly patterns

insights from laboratory-microcosms and land-based aquaculture systems

Doctoral thesis

NTNU
Norwegian University of Science and Technology
Thesis for the Degree of
Philosophiae Doctor
Faculty of Natural Sciences
Department of Biotechnology and Food Science

NTNU
Norwegian University of
Science and Technology

Madeleine S. Gundersen

The effect of disturbances on bacterial community characteristics and assembly patterns

insights from laboratory-microcosms and land-based aquaculture systems

Thesis for the Degree of Philosophiae Doctor

Trondheim, November 2023

Norwegian University of Science and Technology
Faculty of Natural Sciences
Department of Biotechnology and Food Science

NTNU
Norwegian University of
Science and Technology

NTNU
Norwegian University of Science and Technology

Thesis for the Degree of Philosophiae Doctor

Faculty of Natural Sciences
Department of Biotechnology and Food Science

ISBN 978-82-326-7504-3 (printed ver.)
ISBN 978-82-326-7503-6 (electronic ver.)
ISSN 1503-8181 (printed ver.)
ISSN 2703-8084 (online ver.)

Doctoral theses at NTNU, 2023:403

Printed by NTNU Grafisk senter

To 7-year-old Madeleine, who wrote:

"Jeg har lyst til å bli naturforsker".

You did it.

Preface

This research was carried out at the Norwegian University of Science and Technology, at the Department of Biotechnology and Food Science from 2019 to 2023. My PhD position was fully funded by the Faculty of Natural Science. I was affiliated with the research group Analysis and Control of Microbial Systems and was supervised by Professor Olav Vadstein and co-supervised by Professor Ingrid Bakke.

My PhD consisted of the research presented in this thesis in addition to a year of teaching environmental biotechnology (TBT4130), a three-month industrial internship funded by Digital Life Norway and a two-month research stay at the University of Oregon under Professor Brendan Bohannan.

Abstract

Bacterial communities play a vital role in human health, ecosystem stability and various industrial applications, including land-based aquaculture. These communities change over time through predictable deterministic processes such as selection and random stochastic processes such as ecological drift. Our understanding of how disturbances affect community assembly is limited. This knowledge gap hinders accurate predictions of how communities respond to disturbances. This thesis aimed to address this knowledge gap by investigating the effects of various disturbances on community characteristics and assembly processes.

Quantifying community assembly is challenging due to the statistical complexities associated with stochastic processes and their lack of predictable patterns. This thesis first explored the limitations of the null models NTI and βNTI that are used to quantify community assembly. These null models have several assumptions that make them unsuitable for implementation in replicated experimental systems without much environmental variability, such as laboratory systems. Therefore, we propose a novel framework, the 'replicate similarity rate of change', to quantify selection and drift in replicate experimental microcosms with dispersal limitation.

In the context of aquaculture, it has been shown that rearing tank water with a selective pressure for opportunistic bacterial growth reduces the survival of vulnerable fish larvae. However, how the rearing environment affect the characteristics and assembly of the bacterial community in the rearing tank water and in the fish is poorly understood. This thesis investigated this knowledge gap. The results showed that the flow-through, microbially matured and recirculating aquaculture systems influenced the bacterial composition and assembly patterns in the rearing tank water. A key finding was that stochastic community assembly processes in the bacterial communities of the rearing tank water and fish larvae increased over time in a correlated relationship. Furthermore, different rearing environments created distinct stable state attractors, strongly suggesting that microbial management of rearing tank water is feasible. Importantly, we found that small differences in carrying capacity between incoming water and rearing tank water favoured a stable bacterial community with less opportunistic growth and increased larval viability.

Ecosystems such as land-based aquaculture rearing tanks are often disturbed. However, there are many uncontrollable variables in the systems that make it difficult to identify the disturbance mechanisms. Therefore, in the third part of this thesis, we aimed to understand how community characteristics and assembly patterns were affected by the four disturbances periodic dilutions, unsuccessful invasions, antibiotic treatment and phage treatment. In these experiments, bacterial communities from the Trondheimsfjord or Jonsvatnet lake were divided into replicated laboratory microcosms with low environmental variability and operated as semi-continuous systems.

Using the 'replicate similarity rate of change' framework, we found that dilution increased the contribution of selection, while unsuccessful invasions and antibiotics increased drift. The changes in assembly were proposed to be a result of increased resource availability following the disturbance events that initiated opportunistic bacterial growth. Phage treatment had no effect on community assembly, and we found no evidence that the treatment disturbed the communities.

In conclusion, this research shed light on how bacterial community characteristics and assembly processes are affected by disturbances. Furthermore, this thesis provides insights into the management of bacterial communities in aquaculture. Overall, the results pave the way for more accurate predictions of bacterial community responses to disturbances. Such prediction is crucial to ensure the stability of bacterial communities in industrial applications, as well as being relevant to human and animal health.

Acknowledgements

First, I would like to thank my supervisor Olav. It's been a pleasure to work with you. Not only have you pushed me to explore things I thoughts were beyond my limits, but your way of being has been calming, fun and a good match for me. You always remind me that there is more to life than work, and the value in broadening one's perspective. Thank you for giving me the freedom to explore what I found interesting, for giving me so much flexibility and for always taking time to meet with me. I have enjoyed all our small (and long) talks and the friendship we have developed.

I would also like to thank my co-supervisor Ingrid for all advice and input you have given me over the past six years. You have been invaluable in making me learn how to communicate my ideas as our style is somewhat different, which I believe has improved our work. I truly value all the support you both have given me, the responsibilities you have entrusted in me and for creating a safe and including work environment.

My dear colleagues in the AKMS research group. Thank you for making work a fun and pleasant place to be. I appreciate all the times we have overcome challenges together, and mostly with a smile. Kari, you are an inspiration, and although you probably don't know it, you are a role model for me. Ranghild, thank you for giving me the opportunity to join your project and for teaching me all the small tricks on how to maneuver academia (especially the importance of the 9:30 coffee). Amalie, thank you for being a dear colleague, friend and an amazingly caring person who always says yes to coffee. I'm glad I had you on the lab-course team! Eirik, you have a calm presence, and I appreciate all the times you have taken time away from your projects to brainstorm and provide me feedback, and I will miss our daily "God morgen" routine.

Alex, I'm so glad I got to share an office with you throughout my PhD. I lost count of how many fruitful discussions we have had about academia, our papers, our understanding about microbial ecology and life in general. I appreciate how honest (and annoyingly stubborn) you are as a person. With you I really felt like two heads are better than one. Thank you for providing feedback on this thesis, for all the fun we had in Oregon and most importantly for the great academic cooperation we have had throughout the years.

Ian, thank you for letting me join one of your experiments. Although our reactors didn't behave as we planned, it taught me a lot about problem solving, improvising and staying calm. Your love for graphical communication and knowledge in R inspired me to learn this programming language, and maybe spending too much time plotting data. And Tom, thank you for guiding me through some complex modelling problems and teaching me about statistics. It was always a pleasure chatting with you during the lockdown!

Idd and Vetle, my fellow 'foodies' and enjoyers of life. Thank you for all the amazing dinners, wine evenings, rants and laughs we have had. You have been a safe place to vent. Your raw, direct and reflected views on the world reminds me to look beyond the mainstream. You both have a special place in my life.

I would also like to thank my family and friends for supporting and encouraging me. To my dad Andrew, thank you for checking in, helping me when I have needed it and letting me crash your office space during the past couple of months. My mom Anne Karine deserved a special notice for listening to me talk for hours about bacteria, bioreactors, experiments and academia in general, for providing me career guidance and advice, and for inspiring me to be a strong and independent person.

And finally, my dear Jonas. Sometimes I think you understand this thesis better than myself. Thank you for all the times you have listened to my presentations, found solutions to my work-related problems, asked questions about ecology I never thought about, provided me feedback on figures, helped me formulate and communicate what I do, motivated me when everything has seemed pointless, and most of all, for being patient with me every day. You make me feel like I can do anything.

Table of contents

Abbreviations

16S rRNA	Bacterial small subunit ribosomal ribonucleic acid
ASV	Amplicon sequencing variant
βMNTD	Mean nearest phylogenetic neighbour distance
βNTI	Beta nearest taxon index
DGGE	Denaturing gradient gel electrophoresis
DOM	Dissolved organic matter
FTS	Flow-through system
IDH	Intermediate Disturbance Hypothesis
ISH	Intermediate Stochasticity Hypothesis
MMS	Microbially matured system
NGS	Next generation sequencing
OTU	Operational taxonomic unit
PCR	Polymerase chain reaction
QPEN	Quantification of community assembly Processes based on Entire-community Null model analysis
RAS	Recirculating aquaculture system
RC_{Bray}	Modified Raup-Crick index
rrn	Ribosomal RNA

Definitions

General ecology	
Population	The collection of individuals belonging to the same phylotype. A phylotypes can be an ASV, OTU or any other descriptive taxa feature (Kirchman, 2012).
Taxon	A group into which related organisms are classified. For microorganisms, taxa are usually defined by sequence similarity of one or more genomic regions (Hanson *et al.*, 2012). Sometimes used interchangeably with population and species.
Community	The collection of potentially interacting populations that coexist in a defined space at a particular time (Nemergut *et al.*, 2013).
Regional community	The population pool in the entire ecosystem studied. Also known as metacommunity (Hubbell, 2001, Legendre *et al*., 2005)
Local community	The community present in a specific ecosystem under study. The local community is measured in a sample (Legendre *et al.*, 2005).
Community composition	The identity and relative abundance of populations in a sample, assemblage, or community (Hanson *et al.*, 2012).
Mature community	A community that exhibits a stable community composition over time (Odum, 1971).
Pioneer community	The initial community that colonises an ecosystem. It is usually unstable and has low diversity (Odum, 1971).
Microbiota	The collection of microorganisms inhabiting a particular environment such as a human, animal or water (Berg *et al.*, 2020).
Microbiome	The microbiota and their functional properties in a particular environment (Berg *et al.*, 2020).
Resource	All things consumed or used by an organism.
Fitness	An individual's ability to grow and survive in a ecosystem.
Carrying capacity	The maximum population size an ecosystem can sustain given the available resources (Kirchman, 2012).
Niche	A multidimentional space of resources available to and used by a population. The niche characterizes the need a population has for resources, environmental conditions and biotic interactions in order to survive (Hutchinson, 1965; Kirchman, 2012).
Succession	The change in community composition over time. Primary succession describes the changes in community composition occurring during the colonization of a sterile environment (Odum, 1971). Secondary succession describes the community composition changes happening after an environment has been disturbed (Dini-Andreote *et al.*, 2015).
r- and *K*-strategist	A life-strategy framework to categorize organisms based on their maximum growth rate and competitive abilities. *r*-strategists have high maximum growth rates but a low competitive ability, while *K*-strategists have lower maximum growth rates, but higher substrate affinity and utilisation making them better competitors (Andrews & Harris, 1986).
Phage	Viruses that infect bacteria and cause bacterial lysis (Salmond & Fineran, 2015).

	Diversity terms
Diversity	The variety and distribution of unique populations within a particular environment (Tuomisto, 2010, 2011; Shade, 2017).
α-diversity	The diversity within a single local community. α-diversity describes the number of unique populations and their distribution (Legendre *et al.*, 2005).
β-diversity	The difference or change in community composition between two communities (Legendre *et al.*, 2005).
γ-diversity	The diversity in the regional community (Legendre *et al.*, 2005).
Richness	Number of unique populations in a community (Kirchman, 2012).
Evenness	Number of individuals per population in relation to the community (Kirchman, 2012).
	Community assembly
Ecological processes	Mechanisms shaping the composition and diversity within and between communities (Hanson *et al.*, 2012).
Deterministic	Predictable responses. In ecology, deterministic processes include population traits, biological interactions such as competition and mutualism and environmental constraints on populations (Zhou & Ning, 2017).
Stochastic	Unpredictable responses caused by randomness and chance. In microbial ecology drift, extinction, speciation and dispersal are often considered stochastic processes (Zhou & Ning, 2017).
Selection	A deterministic and predictable process caused by differences in population fitness (Nemergut *et al.*, 2013). Similar environmental conditions should lead to homogeneous selection, while different conditions should promote heterogeneous or variable selection (Zhou & Ning, 2017).
Drift	A stochastic process caused by random death resulting in non-predictable fluctuations in population abundances (Nemergut *et al.*, 2013).
Dispersal	Movement of individuals from the regional to the local community (Nemergut *et al.*, 2013). Communities can display dispersal limitation if the movement of individuals is restricted. If the dispersal from the regional to local community is high, dispersal is homogeneous (Zhou & Ning, 2017).
Diversification	Evolution of new genetic types (Nemergut *et al.*, 2013).
Invasion	A special case of dispersal where the individual that moves from the regional to local community constitutes a new population in the local community (Kinnunen *et al.*, 2016).
Propagule pressure	The relative abundance of the invader to the resident community (Kinnunen *et al.*, 2018).
Null models	Algorithms that generate stochastically assembled communities based on properties from an observed metacommunity (Gotelli & Ulrich, 2012).
Null community	The randomly generated community generated by null models.

	Disturbance terms
Disturbance	An event causing changes in the environment or the community. Also known as perturbation (Shade *et al.*, 2012).
Pulse disturbance	A discrete short disturbance event (Shade *et al.*, 2012).
Press disturbance	A long-lasting disturbance event (Shade *et al.*, 2012).
Stable state	The inherent stable composition a community will change towards given the existing composition and environmental conditions. The stable state is also called a locally stable attractor (Scheffer *et al.*, 2001).
Alternative stable state	Following a disturbance, community composition can change to a new stable state.
Resistance	The degree to which the community withstands a disturbance (Shade *et al.*, 2012).
Resilience	Defined as engineering or ecological resilience. Engineering resilience, or rate of return, describes the rate at which the community returns to its original state (Shade *et al.*, 2012). Ecological resilience defines how much disturbance is needed to change the community's properties (Griffiths & Philippot, 2013).
Selective mortality	A mortality event that only affects a particular set of populations in a community (Shade, 2023).
Partial mortality	A mild to moderate mortality event that affects all populations in a community (Shade, 2023).
Mass mortality	A severe mortality event that causes all populations to go, or nearly go, extinct (Shade, 2023).
Stability	The degree to which a community returns to its mean condition after a disturbance and incorporates both resistance and resilience (Shade *et al.*, 2012).
	Other terms
Amplicon	The DNA product after PCR amplification.
Bacteriostatic	Something that causes bacterial growth inhibition (Grenni *et al.*, 2018).
Bactericidal	Something that causes bacterial lysis (Grenni *et al.*, 2018).

List of publications

Paper I. Gundersen, M.S., Morelan, I.A., Andersen, T., Bakke, I., & Vadstein, O. (2021). The effect of periodic disturbances and carrying capacity on the significance of selection and drift in complex bacterial communities. *ISME Communications*, *1*(53). doi: 10.1038/s43705-021-00058-4

Paper II. Vestrum, R.I., Attramadal, K.J.K., Vadstein, O., Gundersen, M.S., & Bakke, I. (2020). Bacterial community assembly in Atlantic cod larvae (*Gadus morhua*): contributions of ecological processes and metacommunity structure. *FEMS Microbiology Ecology*, *96*(9). doi: 10.1093/femsec/fiaa163

Paper III. Gundersen, M.S., Vadstein, O., de Schryver, P., & Attramadal, K.J.K. (2022). Aquaculture rearing systems induce no legacy effects in Atlantic cod larvae or their rearing water bacterial communities. *Scientific Reports, 12:19812*. doi: 10.1038/s41598-022-24149-x

Paper IV. Mathisen, A.J.H, Gómez de la Torre Canny, S., Østensen, M-A, Gundersen, M.S., Olsen, Y. Vadstein, O., & Bakke, I. (in preparation). The early gut microbiome of wild and aquaculture strains of Atlantic salmon is influenced by stochastic processes and environmental bacteria.

Paper V. Gundersen, M.S., Fiedler, A.W., Bakke, I., & Vadstein, O. (submitted). The impact of phage treatment on bacterial community structure is minor compared to antibiotics. Preprint available at Research Square. doi: 10.21203/RS.3.RS-3074836/V1

Paper VI. Fiedler, A.W., Gundersen, M.S., Vo, T.P., Eivind, A. Vadstein, O., & Bakke, I. (2023). Phage therapy minimally affects the water microbiota in an Atlantic salmon (*Salmo salar*) rearing system while still preventing infection. *Scientific Reports 13:19145.* doi:10.1038/s41598-023-44987-7

Work I Gundersen, M.S. (2023, unpublished exploratory work). Exploring pitfalls using NTI and βNTI to estimate community assembly.

Publications not included in this thesis.

Paper VII. Pettersen, J.P., Gundersen, M.S. & Almaas, E. (2021). Robust bacterial co-occurrence community structures are independent of r- and K-selection history. *Scientific Reports 11:23497*. doi: 10.1038/s41598-021-03018-z

List of conference presentations

- Gundersen, M.S., Morelan, I.A., Andersen, T., Bakke, I., & Vadstein, O. **The effect of periodical disturbances on selection and drift in bacterial communities**. Poster presentation; Microbiota and Health, *National Doctoral Programme in Infection and Antibiotics (NPDIA)*, online, 30.11–3.12.2020.

- Gundersen, M.S., Morelan, I.A., Andersen, T., Bakke, I., & Vadstein, O. **Bacterial community assembly during disturbances**. Oral mini-symposium; Contact meeting, *Norsk biokjemisk selskap (NBS)*, Tromsø, Norway, 10-13.02.2022.

- Gundersen, M.S., Morelan, I.A., Andersen, T., Bakke, I., & Vadstein, O. **Disturbances increase selection – whereas stability promotes stochastic assembly in bacterial communities.** Poster presentation; 18th International Symposium on Microbial Ecology, *ISME*, Lausanne, Switzerland, 14-19.08.2022. *Awarded the Bill Costerton award for research that best involves new theories in microbial ecology and wider inter-disciplinary significance.*

- Gundersen, M.S., Fiedler, A.W., Bakke, I., & Vadstein, O. **Antibiotics significantly disturb a drinking water microbiota, while phage treatment has negligible effects.** Poster presentation; 10th IWA Microbial Ecology and Water Engineering Specialist Conference, International Water Association (IWA), Brisbane, Australia, 10-13.10.2023.

“When you can measure what you are speaking about, and express it in numbers, you know something about it; but when you cannot measure it, when you cannot express it in numbers, your knowledge is of a meagre and unsatisfactory kind: it may be the beginning of knowledge, but you have scarcely, in your thoughts, advanced to the stage of *science*, whatever the matter may be” - *Lord Kelvin*

Chapter 1: Bacterial communities - diversity and assembly

Bacteria are everywhere and coexist in communities

It is estimated that there are more bacteria on Earth than stars in the universe. These 10^{30} individuals contain over 15% of the global biomass and have been found in all environments on Earth (Bar-On *et al.*, 2018). Bacteria form complex networks of interactions with each other and their environment (Konopka, 2009). The collective of bacteria living together in a given space at a given time is called a **bacterial community** (Nemergut *et al.*, 2013). These communities can be vastly diverse, consisting of hundreds to thousands of different bacterial types (Konopka, 2009; Kirchman, 2012). They are rich in individuals, and approximately a million bacterial cells are found per millilitre of freshwater and coastal seawater, and around a billion bacteria live in a single gram of soil (Kirchman, 2012). Bacteria not only interact with each other, but also with other inhabitants of the ecosystem, such as viruses, archaea, metazoa and protists. These microorganisms live together in the microbial community and make up the **microbiota** of a particular system (Berg *et al.*, 2020). This thesis examined the bacterial component of the microbiota and will focus primarily on bacterial community research.

Studying bacterial communities in natural ecosystems

Compared to larger organisms, the study of bacteria poses methodological problems because they are too small to be observed with the naked eye. Thus, since Leeuwenhoeck first reported observing microorganisms in his microscope in 1676 , we have relied on technology to observe microorganisms (Leeuwenhoek, 1677).

There are several aspects of a community that can be observed and studied. These range from the description of **community composition** to functional traits. Community composition describes the presence and abundance of each **taxon** (i.e. species, population) (Hanson *et al.*, 2013). The possible ecological roles of different taxa are described by functional traits. As only community composition was studied in this thesis, the analysis of functional traits on community properties will not be described further.

Various methods are used to study the bacterial community, including culture, microscopy and molecular-based approaches (Boughner & Singh, 2016). All approaches have strengths and limitations and can be combined to provide a comprehensive understanding of the bacterial community. Culturing methods rely on isolating and growing bacteria, microscopy uses light or fluorescent emission from individual cells, while molecular methods use DNA- and RNA-based techniques to analyse the genetic material of the bacterial community (Theron & Cloete, 2000; Boughner & Singh, 2016). Culturing methods have provided valuable insights into the presence and distribution of bacteria in various ecosystems (Theron & Cloete, 2000). Unfortunately, typically less than 1% to 10% of the bacterial community is culturable. As a result, culture-based approaches are biased towards culturable populations (Theron & Cloete, 2000). Molecular methods, such as polymerase chain reaction (PCR), denaturing gradient gel electrophoresis (DGGE), and next-generation sequencing (NGS), are now the standard for investigating bacterial communities. Molecular methods provide a more accurate representation of the bacterial community as both culturable and unculturable bacteria are studied.

16S rRNA gene amplicon sequencing – insight at the molecular level

Sequencing of the 16S ribosomal RNA (rRNA) gene is widely used to characterise bacterial communities in different environments. In the late 1980s, Woese identified the 16S rRNA gene as a suitable molecular target for reconstructing evolutionary relationships among bacteria (Woese, 1987). He demonstrated that all bacteria contain conserved regions in the 16S rRNA gene from which universal bacterial PCR primers can be designed. Woese showed that different types of bacteria could be distinguished by examining the regions located between the conserved regions that are prone to mutations and thus vary between populations (Woese, 1987). Over the following decades, advances in DNA sequencing technology and computational methods allowed researchers to rapidly scale up the number of 16S rRNA gene sequences that could be generated, leading to an explosion in the number of studies exploring the diversity of bacteria in different environments.

A general overview of the 16S rRNA amplicon sequencing pipeline consists of DNA extraction, PCR amplification, sequencing and bioinformatics (Figure 1, Jo *et al.,* 2016). The Illumina sequencing platform was used in this work, but other platforms such as Roche 454 and Ion

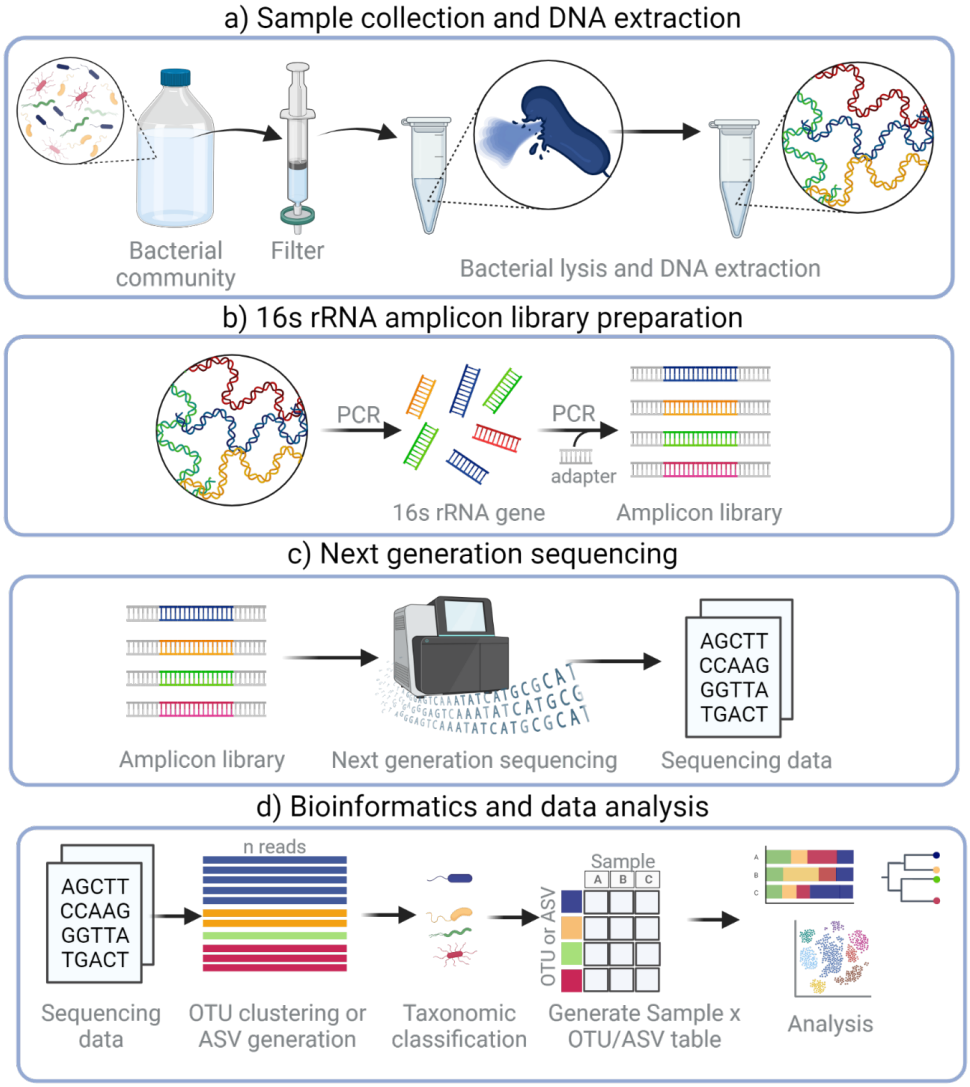


Figure 1: Overview of 16S rRNA amplicon sequencing. a) First, the bacterial community is sampled. For water samples, this usually involves filtering the water at 0.2 µm to collect bacteria on the filter paper. Other sample types require different sampling approaches. Regardless of the sampling method, the next step is to lyse the bacterial cells to extract the bacterial DNA. b) The 16S rRNA gene is then amplified by PCR and, depending on the NGS platform, adapter sequences may be attached in a second round of PCR to barcode the sequences. c) The amplicon library is then sequenced to obtain the genetic information for each amplicon. d) Finally, bioinformatic pipelines generate OTU or ASV tables, which are used to analyse the bacterial community.

Torrent are also available. For Illumina sequencing, bacterial DNA is extracted before the 16S rRNA gene is amplified by PCR using broad coverage primers, resulting in 16S rRNA gene **amplicons** (Figure 1a-b). The PCR amplicons are usually purified before being barcoded with adapters by another round of PCR (Figure 1b). The barcoded amplicons are then sequenced to obtain the genetic information (Figure 1c). Finally, bioinformatics approaches remove low-quality sequencing reads, count the number of reads in each sample, and classify each sequence taxonomically (Figure 1d, Jo *et al*., 2016).

Previously, it was common practice to cluster sequences that were more than 97% similar into an **operational taxonomic unit** (OTU) due to expected sequencing errors and methodological issues. Today, the standard analytical approach is to consider each unique sequence as an **amplicon sequence variant** (ASV) due to the increased precision of sequencing platforms and improved quality control of bioinformatics algorithms (Callahan *et al.*, 2017). The OTUs or ASVs are interpreted as unique populations in the community and are the lowest taxonomic level studied. However, an OTU or ASV do not correspond to a species. A single OTU or ASV may represent multiple bacterial species, in addition to the concept of bacterial species being problematic (see Doolittle, 2012).

Overall, the end product of 16S rRNA gene amplicon sequencing is a count of sequence reads per OTU or ASV from a community sampled in a given space and time. The sequence reads reflect the relative composition of the community and can be used to study bacterial community diversity.

Bacterial community diversity

Bacterial community **diversity** refers to the variety and distribution of bacteria within a particular environment (Tuomisto, 2010, 2011; Shade, 2017). Measuring ecosystem diversity is not straightforward, as there is often no defined boundary from one ecosystem to the next. It is therefore up to the researcher to clearly define the spatial and temporal boundaries of each ecosystem under study.

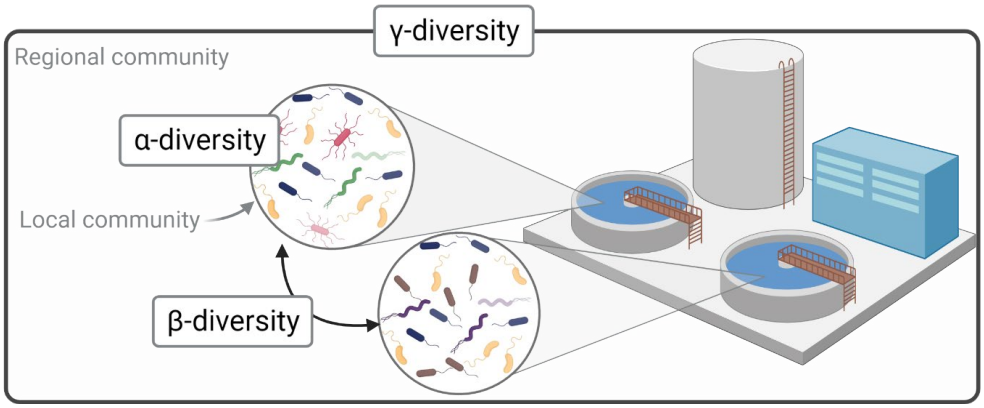


Figure 2: Diversity reflects the variety and distribution of bacteria within a particular environment. For example, in a wastewater treatment plant, γ-diversity describes the diversity of the whole plant, which is the regional community. The diversity in each local community, for example within a treatment tank, is described by α-diversity. Differences between local communities, for example between two tanks, are described by β-diversity.

The regional ecosystem is usually the whole system under study (e.g. a wastewater treatment plant) from which local ecosystems are sampled (e.g. water from a treatment tank). In ecology, α- (alpha), β- (beta) and γ- (gamma) diversity are used to quantify diversity within and between ecosystems (Figure 2, Legendre *et al*., 2005).

α-diversity reflects the number of unique populations (i.e. **richness**) and the distribution of these populations (i.e. **evenness**) in a single **local community** (i.e. sample) (Legendre *et al.*, 2005). Typically, α-diversity is estimated at the OTU or ASV level, but it can be estimated at any taxonomic level. Different α-diversity metrics weight population abundance differently and the choice of metric is a matter of debate. However, a consensus is leaning towards the use of Hill diversity numbers (Lucas *et al.*, 2017). **β-diversity** quantifies the change or difference in composition between two communities. β-diversity can quantify spatial (e.g. between two treatment tanks) or temporal differences in community composition (e.g. same tank at different sampling times) (Legendre *et al.*, 2005). **γ-diversity** incorporates both α- and β-diversity and describes the diversity of the entire **regional community**, also known as the **metacommunity** (Hubbell, 2001; Legendre *et al.*, 2005). The diversity metrics are used to understand and elucidate what is driving the patterns of diversity in the community.

Bacterial community diversity patterns

After the development of 16S rRNA gene amplicon sequencing there was a vast increase in publications characterising bacterial diversity. Comparisons of such studies found that bacterial communities exhibited diversity patterns that were consistent over a range of ecosystems (Nemergut *et al.*, 2013). For example, when populations in an ecosystem are sorted from most to least abundant, almost all bacterial biodiversity studies find a lognormal distribution of population abundance (Figure 3a). Communities are thus highly uneven, with a few dominant taxa and numerous rare ones (Nemergut *et al.*, 2013). Another recurring observation is the positive relationship between bacterial community richness and sample size (Nemergut *et al.*, 2013). That is, metacommunity γ-diversity increases with either increased sampling effort or larger regional ecosystem boundaries (Figure 3b). Furthermore, most bacterial communities exhibit species-time relationships, where richness accumulates over time (Shade *et al*., 2013, Figure 3c). Observing patterns in biodiversity provides insight into ecosystems, but do not explain why these patterns occur. A fundamental goal in ecology is to understand why these patterns arise; *what drives community assembly?*

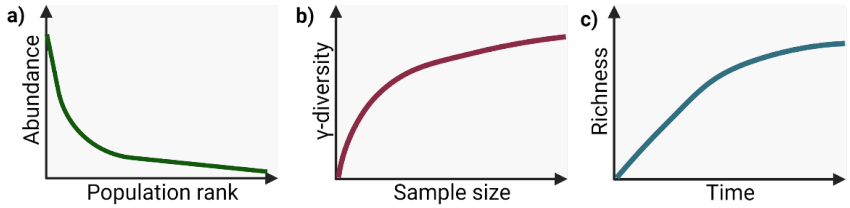


Figure 3: Examples of biodiversity patterns. a) Within a single community, populations generally have a lognormal abundance-population distribution, and when examining the regional community, b) γ-diversity in the regional community tends to increase with increasing sampling effort, and c) population richness accumulates over time.

Bacterial community assembly

Disputes in Community Assembly Theory - Niche vs Neutral Theory

Community assembly theories aim to explain what drives changes in biodiversity patterns. The theories include a set of proposed mechanisms and processes that determine which populations coexist in an area and how they interact. Ecological community assembly theories have been the subject of much debate and research over the years (Wennekes *et al.*, 2012).

These debates have centred on whether **deterministic** (i.e. niche-based) or **stochastic** (i.e. neutral-based) processes are most important in community assembly.

Baas Becking's 'Everything is everywhere, but the environment selects' hypothesis summarises the position of the proponents of niche-based processes (Baas Becking, 1934; Wit & Bouvier, 2006). The hypothesis aimed to explain why similar microbes are found in similar environments. Baas Becking argued that microorganisms are distributed globally by passive dispersal, but that local environmental conditions determine their establishment in a particular habitat (Baas Becking, 1934; Wit & Bouvier, 2006). This notion of environmental restrictiveness was later conceptualised as the niche theory by Hutchinson in the 1950s (Hutchinson, 1965). The **niche** characterises the requirements a population has for **resources**, environmental conditions and biotic interactions in order to survive. Thus, according to niche theory, the available niches in the ecosystem determine the distribution of populations.

MacArthur and Wilson proposed a highly influential theory based on niche theory in their book "The Theory of Island Biogeography" (MacArthur & Wilson, 1967). They proposed that communities are assembled through a process of competitive exclusion. They argued that if species compete for the same niche, all but one species will eventually be driven to extinction. That is, the population with highest **fitness** to the environment would outcompete the rest. The theory suggested that competition could only lead to a stable, co-existing community if species had different niches. Overall, niche-based theories emphasise that community structure is largely shaped by the interactions between populations and their physical and biotic environments.

Niche-based theories are appealing because of their inherent predictability. However, this predictability is the central criticism of the theories, as they fail to explain why there are so many species. If deterministic processes drive community assembly, *should not more species have gone extinct*? This contradiction led ecologists to explore whether neutrality or randomness drives community assembly, and the neutral theory was developed (Rosindell *et al.*, 2011). Neutral theory argues that ecological communities are primarily shaped by stochastic processes such as reproduction, death and dispersal, rather than by species-specific ecological interactions. The neutral perspective gained support after Hubbell wrote

his controversial book "The Unified Neutral Theory of Biodiversity and Biogeography" (Hubbell, 2001). Hubbell's theory assumes that species interaction is negligible for community assembly and that all populations follow the same ecological rules. With these assumptions, he was a pioneer in showing that patterns of biodiversity can be generated by drift and dispersal without including fitness and species-specific interactions. Hubbell argued that stochastic processes such as evolution, immigration and death could produce many of the observed differences between communities.

Neither niche theory nor neutral theory can adequately explain all observed patterns of biodiversity. Thus, a myriad of unique assembly hypotheses were proposed to improve their ability to explain observations (Vellend, 2016). The different models made it difficult to compare studies, as the underlying assumptions and subsequent interpretation of results were based on different hypotheses. To unify the field of community assembly, Vellend argued in “Conceptual synthesis in community ecology” that all community patterns could be explained by four high-level assembly processes (Vellend, 2010). These ideas were elaborated in his book “The Theory of Ecological Communities” (Vellend, 2016).

A unified framework for community assembly

In Vellend's synthesis, both niche-based and neutral processes affect community assembly. Vellend proposed that the underlying mechanisms of community assembly can be attributed to four processes: selection, ecological drift, dispersal and speciation. The relative contribution of these four processes determines how communities change over time and subsequently over spatial scales. Vellend's framework was developed for macrobial communities, and was first adapted to the microbial context by Nemergut *et al.* (2013). Vellend's framework takes a regional-wide perspective of the ecosystem to explain changes in a local community (Figure 4). In his community assembly theory, diversification (or speciation) and dispersal increase biodiversity at the local community scale. In contrast, selection and drift reduce biodiversity over time (Vellend, 2010, 2016).

Selection is a deterministic process and occurs through abiotic and biotic forces. Examples of selective abiotic forces are pH, nutrient availability and salinity. Biotic selection is driven by interactions between two or more individuals. Examples include competition, commensalism

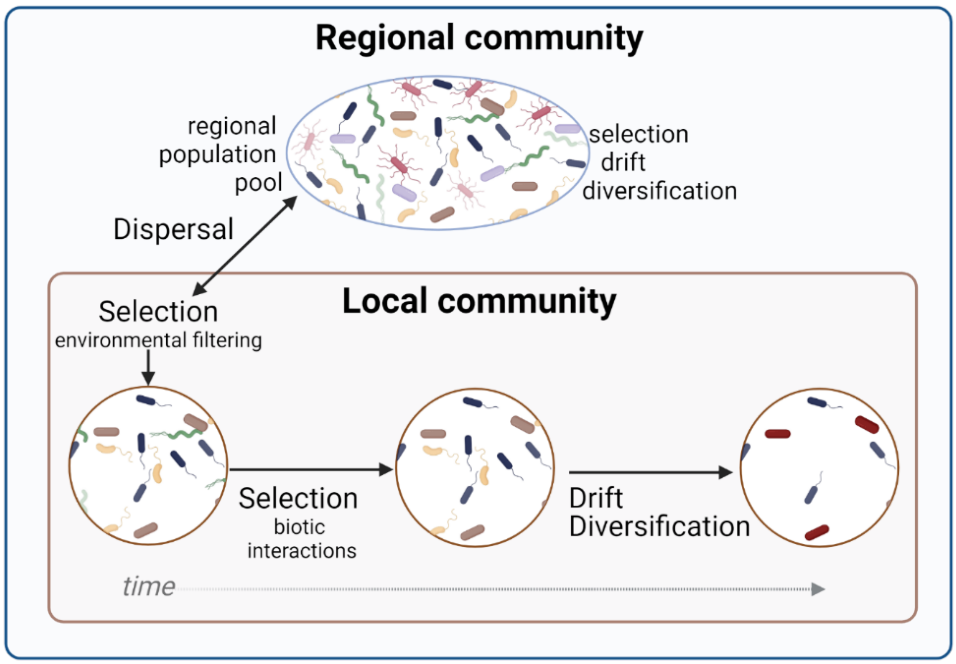


Figure 4: The Theory of Community Assembly proposes that selection, drift, dispersal and diversification shape regional and local communities over time. Dispersal adds individuals to the local community from the regional population pool. Establishment of the population is determined by selection through environmental filtering. Over time, community composition within the local community is shaped by selection (environmental filtering and biotic interactions), drift and diversification. Diversification (i.e. speciation) can create new populations, while selection and drift can reduce the abundance of populations or cause them to go extinct. Figure inspired from Vellend (2010 and 2016).

and mutualism (Nemergut *et al.*, 2013). If selection dominates community assembly, similar environments will produce similar community compositions through homogeneous selection. In contrast, different environments result in different communities through heterogeneous selection (Zhou & Ning, 2017).

In contrast, **drift** is a random process caused by stochastic changes in population abundances (Vellend, 2010, 2016; Nemergut *et al.*, 2013). Until a decade ago, it was debated whether these stochastic changes were relevant in microbial communities due to their large population sizes compared to macrobial communities. However, today it is widely accepted that drift plays a fundamental role in the community assembly of bacterial communities (Zhou & Ning, 2017). The effect of drift is an observable random fluctuation in the abundances of

different taxa (Vellend, 2010, 2016). The most critical consequence of drift is the extinction of low abundance populations (Nemergut *et al.*, 2013).

Dispersal is the movement of an individual from the regional species pool into the local community (Vellend, 2010, 2016; Nemergut *et al.*, 2013). In the case where the population is new to the community, the dispersal event can be classified as an **invasion**. Although some bacteria have evolved mechanisms for directional movement (e.g. chemotaxis and flagella), most bacterial dispersal is a random process because the small size of bacteria allows them to be transported in air or water or attached to other organisms (Zhou & Ning, 2017). Nevertheless, in recent decades, biogeographical patterns have been observed in bacterial distributions, suggesting some deterministic constraints on bacterial dispersal (Hanson *et al.*, 2012). Such a biogeographic constraint is referred to as dispersal limitation (Hanson *et al.*, 2012). When considering smaller regional ecosystems bacteria may exhibit homogenising dispersal, with high rates of dispersal between the regional and local community (Leibold *et al.*, 2004; Zhou & Ning, 2017). Such homogenising dispersal is common in flowing water. Thus, dispersal is influenced by both stochastic and deterministic processes, depending on the environmental conditions and the habitat size under study.

Speciation was extended to **diversification** by Nemergut *et al*. (2013) to account for the evolution of new genetic types. The argument for this change was that, in addition to the ambiguity of the species concept for microbes, the effects of evolution for microbes could have broader community implications than simply the generation of new species (e.g. new metabolic pathways). Diversification is a stochastic process at the molecular level, as mutations occur randomly. However, environmental conditions can deterministically influence the outcome of mutations by increasing selective pressure. For example, toxic compounds such as antibiotics, UV exposure and predation can favour mutations that lead to increased fitness to the conditions. Thus, while diversification is driven by stochastic processes, it can result in deterministic community assembly patterns.

The categorisation of community assembly into selection, drift, dispersal and diversification allowed studies to be compared. Such comparisons has proved to be of great benefit to the field. As a result, there has been a significant increase in the number of papers published on

community assembly (Figure 5). The publication trends reflect the consensus that both stochastic and deterministic assembly processes are essential (Figure 5c). I speculate that it can be a combination of reasons for the significant increase in publications. The main ones being Vellend's "Theory of Ecological Communities", the development of analytical approaches to quantify community assembly and, importantly, the advances in high-throughput sequencing that allow microbial ecologists to characterise community composition relatively well both temporally and spatially (Zhou & Ning, 2017). The latter is clearly reflected in the proportion of assembly studies focusing on microbes, which has increased from 15% to over 50% in the last decade (Figure 5a).

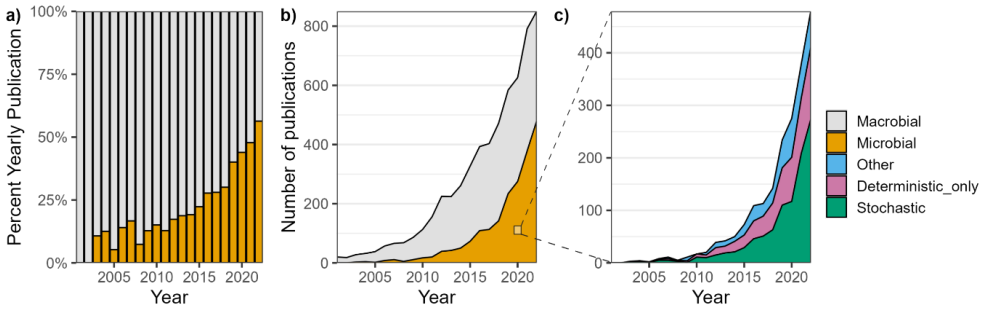


Figure 5: The number of publications on community assembly is increasing. a) The percentage of yearly publications investigating microbial community assembly is rising. b) Studies investigating assembly are steadily increasing for both macrobial and microbial communities, as indicated by the rise in the number of yearly publications. c) Within microbial publications, studies consider stochastic, deterministic only, or other processes to contribute to community assembly. This figure is an extension of the same analysis performed by Zhou & Ning (2017, Figure 1), which included data up to 2016. I used the same search terms as described in their paper on 22.03.2023, but included data from 2000 to 2022.

Quantifying community assembly remains a challenge

While studies on community assembly are increasing, quantifying assembly processes remains a challenge. Communities are highly complex systems and disentangling the magnitude of interacting factors is overwhelming. The main challenge is to quantify stochastic processes, which by definition result in a lack of pattern in community composition. Over the past decades, many different analytical approaches have been developed to infer community assembly processes (e.g. Fukami, 2004; Legendre *et al.*, 2005; Sloan *et al.*, 2006; Chase & Myers, 2011; Stegen *et al.*, 2012; Ovaskainen *et al.*, 2017; Ning *et al.*, 2019, 2020). Approaches that aim to disentangle deterministic and stochastic processes can be classified into three

main categories: multivariate analysis, neutral models, and null models (Vellend *et al.*, 2014; Zhou & Ning, 2017).

In brief, multivariate approaches aim to explain variation in the dataset based on the measured environmental conditions and similarities between the measured communities (Legendre *et al.*, 2005; Vellend *et al.*, 2014). In this approach, the variation explained by the conditions is generally categorised as deterministic processes, while the unexplained variance is considered stochastic. However, a drawback of these approaches is that unexplained variance also includes unmeasured environmental conditions (Clark *et al.*, 2007; Zhou & Ning, 2017).

Neutral models, on the other hand, are mathematical models that contain theoretical parameters used to describe a stochastically assembled community (e.g. Hubbell & Hubbell, 2005; Sloan *et al.,* 2006). Observed data are fitted to the model to estimate the parameters, and the fit of the model is used to determine the degree of stochastic or neutral community assembly. The main criticism of neutral models is that they oversimplify community properties, making them difficult to apply universally (Zhou & Ning, 2017). Furthermore, due to technical limitations, we generally do not have a complete characterisation of bacterial community composition. Low abundance populations are often overlooked, which may be very important in the stochastic processes (Zhou & Ning, 2017).

The **null model** approach was developed to overcome the oversimplification of neutral models. In null model analytical approaches, randomly assembled communities are generated based on the properties of the observed metacommunity (Gotelli & Ulrich, 2012). There are numerous approaches to performing the randomisations, such as keeping community properties like population richness and community size constant. The randomisations are usually repeated many times, and each randomisations represents a **null community**. After generating the null communities, the observed community is compared to the null communities using various methods. Typically, a two standard deviation between the observed community and the null community is considered to be significant. Significant deviations indicates that deterministic processes structured the assembly (e.g. Stegen *et al.,* 2012). However, null model approaches are prone to overestimating stochasticity by

randomising with too few constrains, resulting in an excessively random null community. Similarly, stochasticity can be underestimated by randomising with too many constraints, resulting in too similar null and observed communities (Gotelli & Ulrich, 2012). Thus, the choice of null model and analytical approach will strongly influence the results of the null model analysis.

A caveat with null models is that they only distinguish between deterministic and stochastic processes. It is desirable to further disentangle the community assembly processes. The analytical tool Quantification of community assembly processes based on Entire-community Null model analysis (QPEN) can distinguish selection and dispersal from other stochastic processes (Stegen *et al.*, 2013). QPEN combines the two null models, the beta-nearest taxon index (βNTI) and the abundance-extended modified Raup-Crick index (RC_{Bray}). By combining these models QPEN incorporates both phylogenetic and compositional community properties (Figure 6).

The first step of QPEN is quantifying βNTI (Stegen *et al.*, 2012, 2013). First, the β-mean nearest phylogenetic neighbour distance (βMNTD) is calculated for the observed community comparison and each randomly generated null community comparisons (usually 1000 randomisations). Then, βNTI is quantified as the standard deviation of βMNTD in the observed community comparison from the distribution of βMNTD in the null communities. As such, the βNTI reflects the difference in phylogenetic similarity between two communities. The magnitude of the standard deviation is used to disentangle selection from stochastic processes.

The underlying assumption of βNTI is that taxa with a close phylogenetic relationship occupy the same habitat. Thus, communities containing populations with close phylogenetic relationships are assumed to be assembled by selection processes such as environmental filtering and biotic interactions. Community comparisons with a βNTI<-2 are assumed to have been structured by homogeneous selection, as the communities are more phylogenetically clustered than expected by chance. In contrast, when βNTI>2, the compared communities are more phylogenetically dispersed than expected by chance, indicating that environmental conditions were variable and that heterogeneous selection structured assembly. Community

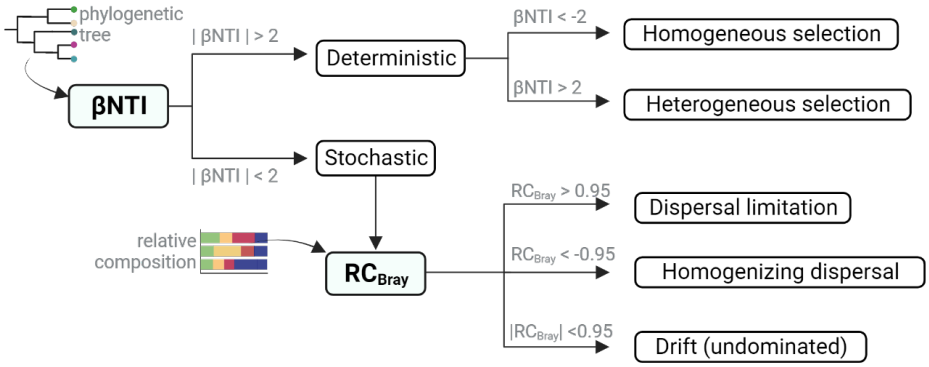


Figure 6: In the QPEN framework, phylogenetic and compositional community properties are used to disentangle selection, dispersal and drift. First, the metacommunity phylogenetic tree is used to estimate βNTI to determine whether two communities are more phylogenetically clustered or dispersed than expected by chance. Comparisons with a |βNTI|>2 are considered deterministic and further classified as structured by homogeneous or heterogeneous selection. In contrast, |βNTI|<2 are assembled stochastically, and RC_{Bray} is used to disentangle dispersal and drift based on compositional data. Comparisons where both βNTI and RC_{Bray} are insignificant are determined to be structured primarily by drift. Drift is sometimes referred to as undominated because it can also be attributed to weak selection, weak dispersal and diversification. Figure adapted from Zhou & Ning (2017).

comparisons that do not deviate from the null model have a βNTI between -2 and 2. These communities are assumed to be structured by stochastic processes. The stochastic processes are further disentangled using RC_{Bray}.

RC_{Bray} is a probabilistic null model based on taxonomic abundance data that tests whether two communities are more or less similar in composition than expected by chance (Chase & Myers, 2011; Stegen *et al.*, 2013). Homogenising dispersal is expected to increase similarity as populations move freely between the two communities. RC_{Bray}<-0.95 indicates that communities are significantly more similar than expected by chance and are assumed to be structured by homogenising dispersal. In contrast, comparisons with RC_{Bray}>0.95 are less similar than expected by chance and can occur when dispersal limitation is strong. Comparisons that are not significantly different from the null model (i.e. -0.95<RC_{Bray} <0.95) are often interpreted as drift. The term undominated is also used as a label for not significant RC_{Bray} estimates as these patterns can occur due to weak selection, weak dispersal and diversification. Thus, while the QPEN framework manages to disentangle selection and

dispersal from other assembly processes, it remains difficult to disentangle drift and diversification (Zhou & Ning, 2017).

QPEN gained popularity because it was the first analytical tool capable of disentangling some of the ecological processes. It has been widely used to quantify bacterial community assembly processes in a variety of ecosystems, such as agricultural soils (Liu *et al.*, 2021), aquifers (Fillinger *et al.*, 2019), coastal waters (Wang *et al.*, 2020) and marine fish larvae (Yan *et al.*, 2016). However, the method makes many assumptions that may not be appropriate for all experimental approaches. First, QPEN assumes that closely related populations exhibit similar ecological characteristics which can be measured as a phylogenetic signal. There is debate about whether such signals exist and, if so, how they should be measured (Zhou & Ning, 2017). While Stegen *et al.* (2012) explicitly state that the presence of a phylogenetic signal should be confirmed, many publications do not report testing this initial assumption (e.g. Jiao *et al.,* 2019; Shi *et al.,* 2018). Second, quantifying selection based on phylogeny is controversial because it assumes that deterministic processes are better reflected in phylogenetic than compositional data (Zhou & Ning, 2017). As a result, quantification of selection is highly dependent on the construction of the phylogenetic tree as well as the size and choice of the metacommunity.

Furthermore, the core underlying assumption of QPEN is that environmental variation will select for communities that are more or less phylogenetically similar than expected by chance. This assumption of environmental variability is often not met in controlled experimental laboratory conditions where variability often is minimised by design. Thus, there is a need to develop analytical approaches that can quantify community assembly in controlled experimental laboratory conditions.

Although tools such as QPEN have limitations, they have been invaluable in accelerating our understanding of community assembly processes in bacterial communities. Such knowledge is essential for understanding how disturbances affect community assembly processes.

Chapter 2: Stability and disturbances in bacterial communities

Chapter 1 introduced the concept that populations assemble into communities through ecological community assembly processes. Chapter 2 introduces key concepts related to community response to disturbances and aims to present the link between knowledge of disturbances and community assembly.

Disturbances disrupt stability in community composition

Disturbance initiates secondary succession

Communities change over time through **succession** (Odum, 1971)**.** Traditionally, succession has been defined as a niche-based process in which two identical but separate communities develop and assemble similarly over time under the same environmental conditions (Zhou *et al.*, 2014). The initial colonisation of an uninhabited environment is referred to as primary succession and is usually occupied by a **pioneer community** (Odum, 1971). Members of the pioneer community interact and stability in their abundance distributions is achieved over time. The community then transitions to a **mature community** where community composition is stable (Odum, 1971). Such ecological **stability** is largely a theoretical concept. It is normal for ecosystems to experience disturbances that disrupt stability.

Disturbances are events that alter either the physicochemical environment or the community (Shade *et al.*, 2012). Depending on how the community responds to the disturbance, community composition may or may not change (Shade *et al.*, 2012). Disturbances can initiate secondary succession, in which communities reassemble in the previously colonised ecosystem (Dini-Andreote *et al.*, 2015). Depending on the type and severity of the disturbance, community members may not respond, change in abundance or go extinct (Shade *et al.*, 2012).

Disturbances are classified as pulse or press disturbances

Depending on how long the disturbance event persists one can classify disturbances as pulse- or press- disturbances (Shade *et al.*, 2012). **Pulse disturbances** are generally short and discrete events such as a wildfire (Ferrenberg *et al.*, 2013), a drought (Chase, 2007), or a storm

(Herren *et al.*, 2016). Because pulses are discrete events, the community can respond to the disturbance and recover to its original state (Shade *et al.*, 2012). **Press disturbances** are, on the other hand, persistent and result in lasting changes in the environment (Shade *et al.*, 2012). Examples include increased organic loading in a treatment reactor (Santillan *et al.*, 2020) or increased temperature (Sorensen & Shade, 2020). These changes require the community to adapt to the changes.

Community response to disturbances – resistance and resilience

Response is defined by community resistance and resilience

Different terms are used to describe the response of a community to disturbance. **Resistance** is the degree to which the community withstands the disturbance, while **resilience** can be defined as engineering or ecological resilience (Shade *et al.*, 2012; Griffiths & Philippot, 2013). Engineering resilience, or rate of return, describes the rate at which the community returns to its original state (Shade *et al.*, 2012). Ecological resilience defines how much disturbance is required to change the characteristics of the community (Griffiths & Philippot, 2013). In general, engineering resilience decreases as ecological resilience decreases; that is, the rate of return decreases as the disturbance required to introduce change decreases (Van Nes & Scheffer, 2007). The degree of adaptability or response of a community to disturbance is referred to as **stability**. Stability is the degree to which a community returns to its mean state following a disturbance and includes both resistance and resilience (Shade *et al.*, 2012). The response of a community to a disturbance depends on the habitat conditions, the type and duration of the disturbance and community properties such as composition, function and growth strategy (Philippot *et al.*, 2021).

Disturbances can disrupt a community from its stable state

Stability is a stable state that communities tend to change towards. Community composition can be envisioned as a marble that rolls towards the lowest (i.e. most stable) point in an ecosystem landscape (May, 1974, Figure 7a). Scheffer *et al*. (2001) argued that a communities point of stability, or inherent **stable state,** is given by the community composition and environmental conditions. This inherent stable state is also known as a **locally stable attractor** because the community continuously change towards the stable state composition. The communities do not necessarily recover to their pre-disturbance composition after a

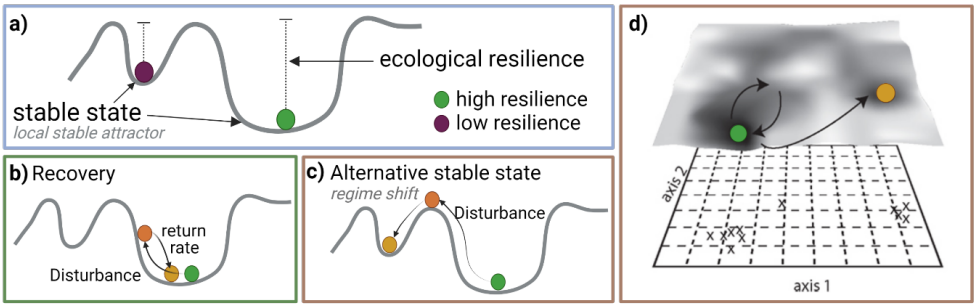


Figure 7: Communities obtain a stable state given their environmental conditions, which can be imagined as valleys in an ecosystem landscape. a) These valleys or locally stable attractors define the ecological resilience of the communities. b) Disturbances can change the community composition. Depending on the degree of disturbance, the community may recover to its original community composition. c) Strong disturbances or low resilience can cause a regime shift, where the community transitions to a new alternative stable state. d) The stability of a community can be visualised in an ordination plot. Stable states can be detected where samples cluster in an ordination space. Regime shifts can be observed if samples cluster in a different ordination space following a disturbance. Figures a-c were inspired by Shade *et al.* (2012) and Van Nes & Scheffer (2007), and d was modified from Shade *et al.* (2012).

disturbance. Disturbances can cause shifts in community composition leading to a **new alternative stable state** if ecological resilience is low or the disturbance is substantial (Scheffer *et al*., 2001; Shade *et al.,* 2012, Figure 7b-d).

How do bacterial communities respond to disturbances?

Understanding how bacterial communities respond to disturbance is central to the study of ecosystem diversity. A puzzling observation in nature is the unexpectedly high diversity found in resource-limited homogeneous environments, which defies the expectation that competitive exclusion gradually reduces biodiversity over time. This ecological phenomenon, known as the 'Paradox of the Plankton', was originally explained by the formation of niche differentiation (Hutchinson, 1961). However, Paine offered an alternative explanation and linked the Intermediate Disturbance Hypothesis (IDH) to the 'Paradox of the Plankton' by proposing that intermediate disturbances lead to increased diversity (Paine, 1969).

The Intermediate Disturbance Hypothesis is related to stochasticity

IDH proposes that diversity in ecosystems is maximised at intermediate levels of disturbance (Connell, 1978). The idea is that disturbance alters the occupancy and types of niches available in an environment. Intermediate levels of disturbance, thus, create opportunities for new species to colonise available niches, but prevent competitive exclusion (Osman, 2015). This hypothesis suggests that ecosystems that experience too little disturbance may become dominated by a few highly competitive species. In contrast, ecosystems that experience too much disturbance may become unstable and prone to frequent extinction events. The IDH has been empirically supported in a wide range of ecosystems (Osman, 2015). However, its validity has been debated due to the lack of universal observational support for increased diversity at intermediate disturbances (Fox, 2013, 2013; Sheil & Burslem, 2013).

Nevertheless, IDH is a recognised framework for explaining patterns of diversity and has been linked to community assembly processes. Santillan *et al.* (2019) disturbed replicate activated sludge microcosms with a toxin at various frequencies, ranging from never to daily pulses over 35 days (Santillan *et al.*, 2019). Based on their observations, they proposed the Intermediate Stochasticity Hypothesis (ISH), which predicts that community biodiversity is highest at intermediate disturbances due to increased stochastic assembly following the disturbance. They argued that predictable environments, such as undisturbed or daily disturbed environments, increase the fitness of populations with specialised traits. Thus, at extreme levels of disturbance, determinism increases because competitive advantages are high, leading to competitive exclusion and subsequently reduced α-diversity. In unpredictable environments, such as those with intermediate levels of disturbance, this competitive advantage decreases. The lack of competitive advantages is then expected to increase the contribution of stochastic processes and increase α-diversity. Santillan & Wuertz (2022) found support for ISH using nutrient input as a disturbance. Thus, ISH appears valid in response to toxin and nutrient disturbances. ISH demonstrates that assessing community assembly patterns following disturbance can improve our understanding of community responses to disturbance which currently is poor.

Community responses to specific disturbances

Over the past decade, there has been an increased focus on understanding how disturbances affect community assembly. However, there is a significant gap in our understanding of the underlying ecological mechanisms. The term 'disturbance' is broad as it encompasses any change to an ecosystem. Therefore, this introduction will focus on the knowledge gaps regarding the disturbances applied in this thesis. I have focused on the current state of knowledge of how these disturbances affect community diversity, successional patterns and community assembly. Although disturbances can affect community functionality (Yachi & Loreau, 1999), this is beyond the scope of this thesis. The disturbances considered in this thesis can be divided into the following categories: increased resource availability, increased mortality and bacterial invasions. These categories are discussed in more detail below.

Increased resource availability

The resources available to a community are fundamental to what types of bacteria that coexist as it determines the **carrying capacity** of an ecosystem (Kirchman, 2012). The carrying capacity is the maximum population size an ecosystem can sustain over time given the available resources. Disturbances in resource availability can, therefore, drastically change community composition. Resource availability per individual can be increased by adding resources to a community or by removing community members. The addition of resources effectively increases the carrying capacity of the community, while the latter reduces biomass below carrying capacity. Biomass reduction is further discussed under mortality in the following section.

To what degree increasing resource availability alters the community assembly patterns has been studied to some extent during the past decade. Zhou *et al.* (2014) developed two hypotheses regarding how nutrient input disturbances should affect community assembly (Figure 8). In brief, they argued that stochasticity increases in communities with high dispersal due to the growth of the rare biosphere, weakened niche selection pressure and strengthened priority effects. According to their framework, such highly connected systems will eventually recover when environmental conditions return to the pre-disturbed state (Figure 8a). In contrast, dispersal-limited communities would have the same initial responses as those with high dispersal but should not recover as environmental conditions have shifted (Figure 8b, Zhou *et al.,* 2014).

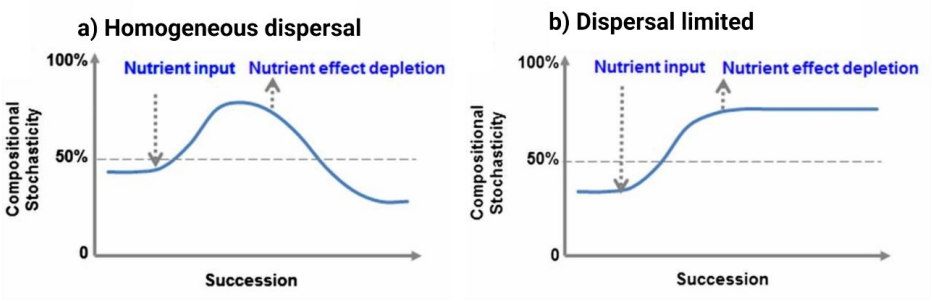


Figure 8: Stochasticity is influenced by resource increases on short ecological time scales according to Zhou *et al.* (2014). Dispersal from a regional population pool to local communities can be homogeneous or limited, depending on ecosystem connectivity. a) Communities with homogeneous dispersal are generally highly connected to the regional ecosystem. In highly connected communities, stochasticity is expected to increase as resources increase. However, due to high dispersal rates and diffusion of resources into the regional ecosystem, the community recovers and stochasticity decreases. b) In contrast, dispersal-limited communities do not have this connectivity and have a harder time recovering as resource conditions remain altered. Thus, stochasticity increases after resource increases and remains high. Figure adapted from Zhou *et al.* (2014).

There is some support for their hypothesis. Zhou *et al*. (2014) tested their hypothesis in a system with homogeneous dispersal and found that stochasticity increased following the nutrient input and thereafter gradually decreased. In a dispersal-limited community, Santillan *et al.* (2020) observed that pulses of organic loading increased the contribution of stochastic processes in sequencing batch reactors. While it is predicted that stochasticity increases due to perturbations in nutrient enrichment, more research is needed to validate this hypothesis, as few studies have examined the relationship between resource enrichment and community assembly.

Furthermore, there is evidence that different populations have different life-strategic responses to increases in resources (Klappenbach *et al.*, 2000; De Vries & Shade, 2013). Bacteria can be classified as ***r*- or *K*-strategists** based on their maximum growth rate and substrate utilisation ability (Andrews & Harris, 1986). *r*-strategists thrive under environmental conditions of excess resource availability because they have high maximum growth rates but are poor competitors. In contrast, *K*-strategists, are highly competitive due to higher substrate affinity and utilisation constants. Thus, *K*-strategists thrive when biomass is close to or at the carrying capacity of the ecosystem and competition for resources is strong (Andrews & Harris, 1986).

Studies have shown that adding nutrients to seawater results in a bloom of fast-growing bacteria, indicating that *r*-strategic bacteria exploit the excess resource situation. The response to nutrient increases was detectable at the phylum level, where the relative abundance of *γ-proteobacteria* increased drastically following nutrient additions, while *α-, β- and δ-proteobacteria, Flavobacteria* and *Verrucomicrobia* had higher abundances when the resource landscape was more stable (Vadstein *et al.*, 2018). These results suggest that the *r/K*-life strategies are evolutionarily adapted and, as such, can be detected at the molecular level. Indeed, it has been observed that *r*-strategic bacteria contain more ribosomal RNA (rrn) operons than *K*-strategic bacteria (Klappenbach *et al.*, 2000; Roller *et al.*, 2016). Averaging the number of rrn operons for all populations in a community can give insight into the life strategy at the community level. It has been shown that the average number of rrn operons in a community increases after resource increases (Kearns & Shade, 2018). These observations indicate that *r*-strategists are the first responders to resource increases, which is expected given their ability to exploit environments with excess resources.

The rapid response of *r*-strategists to resource increase, and the consequent consumption of excess resources, has been suggested to shorten the recovery time of communities following resource increases (De Vries & Shade, 2013). Indeed, the recovery time following a disturbance was faster in soils with high than with low resource availability and was associated with an increased abundance of *r*-strategists (Frenk *et al.*, 2018). While the growth response of *r*-strategists is deterministic, the overall community assembly pattern may be stochastic if many types of *r*-strategists can respond to the disturbance. Thus, investigating the relationship between community assembly and *r*/*K*-strategist responses following disturbance emerges as an exciting framework for understanding disturbance responses.

Two important research areas need to be addressed regarding increased resource availability. Firstly, it is important to validate if stochasticity increases due to nutrient enrichment to deepen our understanding of the relationship between community assembly and increased resource availability. Secondly, it is valuable to relate community assembly changes to the *r/K*-life strategy to better understand which types of bacteria that respond to resource disturbances.

Selective, partial and mass mortality following disturbance

Some disturbances result in reduced population densities or loss of biomass. Disturbances can result in selective, partial or mass mortality at the community level (Shade, 2023). **Selective mortality** occurs when the disturbance affects only a particular set of populations and can be density-dependent or -independent. Examples of disturbances that induce selective mortality are size-selective predation by flagellates (Hahn & Höfle, 1999), antibiotics (Grenni *et al.*, 2018) or viral predation by bacteriophages (Johnke *et al.*, 2014). **Partial-** and **mass mortality** occur when all populations, regardless of density, are affected by the disturbance. Examples of disturbances that cause partial mortality are dilutions (Gibbons *et al.*, 2016) and unselective predation (Johnke *et al.*, 2014). Mass mortality occurs when the disturbance causes all populations to go or nearly go extinct (Shade, 2023). Events that cause mass mortality can be severe radiation exposure (Pullerits *et al.*, 2020), prolonged heat treatments (Coma *et al.*, 2009) or heavy chemical spills (Low-Décarie *et al.*, 2015). The loss of individuals from a community makes niches available, and the interaction network in the bacterial community is disrupted and changes (Philippot *et al.*, 2021). Thus, disturbances that cause biomass loss can alter the diversity and composition of bacterial communities.

Dilution: density-independent partial and mass mortality

Dilution is a density-independent disturbance because all populations are reduced by the dilution factor. Dilutions often occur with solutions containing resources. Thus, when community size is reduced following such dilutions, the resources available per individual increase, even though carrying capacity remains unchanged. Such dilutions have two effects: biomass loss and increased resource availability.

According to Zhou *et al.* (2014), disturbances that result in biomass loss are expected to increase the contribution of deterministic processes mainly by enhancing niche-selection. In contrast, disturbances that enhance resource availability are expected to increase the contribution of stochastic processes due to weakened niche selection (Figure 8, Zhou *et al.,* 2014). Furthermore, at the local community level, low abundance populations are vulnerable to dilution because they have a high probability of being diluted to extinction (Zhang & Zhang, 2015). Although such extinction is a stochastic process, it can lead to a deterministic pattern over time if only the resilient populations with higher growth rates persist. Thus, from a

theoretical perspective, both stochastic and deterministic assembly can be enhanced by dilution. It is currently not well understood if deterministic or stochastic assembly dominates following a dilution.

Some studies have aimed to investigate the relationship between dilutions and community assembly. Gibbons *et al*. (2016) disturbed aquatic lake-derived communities through a range of dilution strengths and frequencies. Interestingly, they observed increased diversity and between-sample variation at intermediate levels of dilution. At the same time, the intermediately diluted communities were the most dissimilar to the control communities. Thus, stochastic processes were most pronounced at intermediate levels dilution, supporting IDH and ISH. In contrast, Mao *et al.* (2023) found no evidence for IDH or ISH in a dilution-to-extinction experiment. Instead, stronger dilution strengths increased the contribution of deterministic processes, reduced diversity, and increased the average rrn operons in the community. This shift in community composition indicated that *r*-strategic bacteria were selected for when resources became available (Zhang & Zhang, 2015). Community assembly responses to dilutions are thus varying.

It is not clear how community assembly processes are affected by dilution. The dilution factor, frequency and observation window are likely to affect the community assembly processes. Investigations into community assembly responses to dilutions are therefore needed.

Antibiotics: broad-targeting selective mortality

Antibiotics are molecules that prevent vital metabolic pathways in bacteria, resulting in growth inhibition (**bacteriostatic**) or bacterial lysis (**bactericidal**) (Grenni *et al.*, 2018). Antibiotics can cause significant changes in the diversity, functionality, stability and composition of bacterial communities (Yassour *et al.*, 2016; Grenni *et al.*, 2018; Cairns *et al.*, 2020; Bent *et al.*, 2021). In this thesis, oxytetracycline and Penicillin G were used as antibiotic disturbances. Oxytetracycline (5-hydroxytetracycline) is a tetracycline that inhibits protein synthesis by blocking tRNA from binding to the ribosomal unit in both Gram-negative and Gram-positive bacteria (Chopra & Roberts, 2001). As a result, the bacteria are unable to divide due to a halt in metabolism. The tetracycline binding is reversible, and therefore, the antibiotic is classified as a bacteriostatic compound (Chopra & Roberts, 2001). On the other

hand, the β-lactam Penicillin G (benzylpenicillin) is a bactericidal antibiotic. Penicillin G affects Gram-negative and Gram-positive bacteria by inhibiting enzymes involved in cell wall synthesis causing lysis of the cell membrane (Bush & Bradford, 2016). The two antibiotic compounds have different modes of action; oxytetracycline results in growth inhibition, whereas Penicillin G causes bacterial lysis.

Antibiotics are thought to enhance deterministic community assembly as they act as a selective filter against targeted bacterial populations (Antonopoulos *et al.*, 2009; Grenni *et al.*, 2018). Most ecological studies of community responses to antibiotic exposure have focused on how antibiotic resistance genes are selected for. Very few studies have examined changes in community assembly in response to antibiotic exposure.

Recently, two studies reported that antibiotic treatment increased the contribution of stochastic processes (Cairns *et al.*, 2020; Chen *et al.*, 2023). Cairns *et al.* (2020) found that increasing concentrations of the antibiotic streptomycin increased stochasticity by driving many populations to extinction in a 34-species synthetic community. Higher concentrations of streptomycin had a more substantial negative impact on diversity and made it more challenging for the community to recover to its pre-disturbed state. In a more complex soil derived community, Chen *et al.* (2023) found that a one-time exposure to a mix of eleven different antibiotics enhanced the contribution of drift. If such stochastic patterns are a universal response to antibiotic exposure is unknown and our knowledge of the effects of antibiotics on community assembly is limited. Investigations using different types of antibiotics and other ecosystems should be performed to validate these observations.

Phages: Strain selective mortality

Bacteriophages, or **phages**, are viruses that infect bacteria (Salmond & Fineran, 2015). They are the most abundant and diverse biological entities on the planet and have been found in every ecosystem where bacteria exist (Dion *et al.*, 2020). Phages were first discovered by Twort in 1915 and independently by d'Hérelle in 1917 (Chanishvili, 2012). Twort coined the term bacteriophage, meaning bacteria eater, as he observed that something extremely small made turbid bacterial cultures clear (Chanishvili, 2012). It was later discovered that this disappearance of bacteria was caused by lysis of the bacterial cell membrane during the final stage of the phage's life cycle (Figure 9).

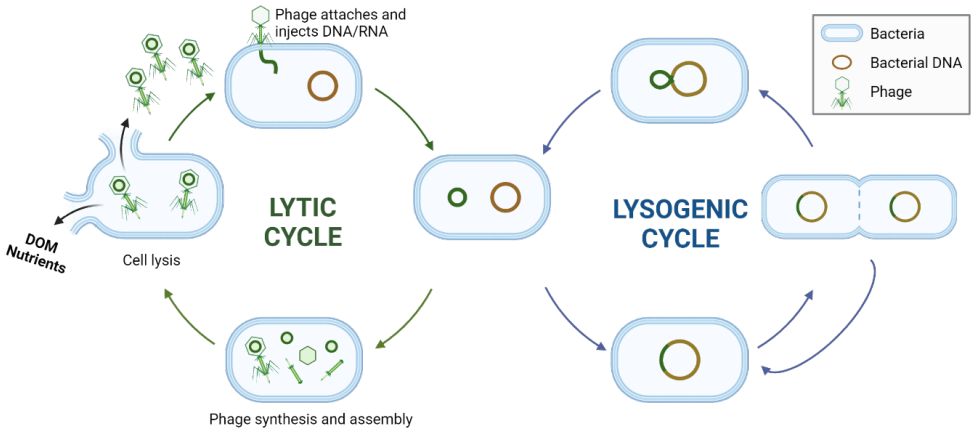


Figure 9: Phages can have two life cycles - lytic or lysogenic. First, the phage binds to the bacterial host and injects its genome. Virulent phages enter the lytic cycle, while temperate phages can follow either the lytic or lysogenic cycle. In the lytic cycle, the phage hijacks the bacterial replication system to make new viral particles. Finally, the cell is lysed, releasing the new viral particles, DOM and nutrients into the environment. The burst size of the viruses from the bacteria ranges from a few to hundreds of new viral particles, although the average is around 50 (Kirchman, 2012). In the lysogenic pathway, the viral genome is integrated into the bacterial DNA. The lysogenic cycle usually transitions to the lytic cycle when the bacterium becomes stressed (Salmond & Fineran, 2015).

Phages can be used as therapeutic agents to remove unwanted bacteria from an ecosystem because of their ability to lyse bacterial cells. This approach, known as phage therapy, was successfully used to treat dysentery as early as 1919, but development of phage therapy was halted due to the discovery of antibiotics (Chanishvili, 2012). Today, there is increased interest in expanding our knowledge and development of phage therapy due to the increasing resistance bacteria have to antibiotics, which is one of the top ten global health threats (WHO, 2022).

Phages have several properties that make them suitable candidates to replace antibiotics. For example, phages are bactericidal rather than bacteriostatic and multiply through the lytic cycle when the bacterial host is present. Furthermore, phages have a narrow host specificity, even down to the strain level (Ganeshan & Hosseinidoust, 2019). These properties are advantageous for the treatment of infections, as the specific pathogen can be targeted and the need for repeated treatment is reduced. However, phages and bacteria continuously

evolve to gain a competitive advantage in their predator-prey relationship (Thingstad, 2000). This co-evolution could pose a risk, as the phage treatment could become less effective if the target bacteria develop resistance. Furthermore, because phages are highly specific, they need to be isolated and characterised for new infections, which can take days to weeks (Gill & Hyman, 2010).

There is a lack of knowledge about the effects of phage therapy on the resident community (i.e. the untargeted bacterial community). From an ecological perspective, phage therapy can act as a disturbance by increasing available resources and altering the bacterial interaction network. For example, in natural ecosystems, phages are responsible for 5-50% of bacterial mortality and thus play a crucial role in maintaining bacterial productivity and turnover (Proctor & Fuhrman, 1990; Kirchman, 2012). When phages induce lysis of target bacteria, micronutrients and dissolved organic matter (DOM) are released into the environment, which can be used for growth by the resident community (Hess-Erga *et al.*, 2010; Kirchman, 2012; Vadstein *et al.*, 2018).

Furthermore, established microbial communities have complex networks of interactions. Thus, removal of one population may have a cascading effect in the network, altering competitive or mutualistic relationships. To my knowledge, only one study has investigated the effect of disrupting interaction networks using phages. Wu *et al.* (2022) investigated this potential cascade effect by inoculating germ-free mice with a 10-species synthetic bacterial community and removing two populations at day 16 and two others at day 30 using phages. They found that the removal of populations did indeed alter the interaction network (Wu *et al.*, 2022). Thus, phage therapy can disturb bacterial communities.

There are few studies investigating the effect of phage therapy on changes in bacterial communities. Most studies have assessed if it is safe for an organism to consume phages and have been conducted in the absence of the phage host (i.e. uninfected individuals). The majority of these safety-focused studies found no effect of phage therapy on bacterial communities in humans (McCallin *et al.*, 2013; Sarker *et al.*, 2017; Grubb *et al.*, 2020), chickens (Juan *et al.*, 2022) and turtles (Ahasan *et al.*, 2019). However, it has been shown in chickens that phage therapy induced changes in the bacterial community even in the absence of the

phage host (Zhao *et al.*, 2022). When the host is present, more studies find an effect of phage therapy, for example in rainbow trout (Donati *et al.*, 2022), synthetic human gut communities (Hu *et al.*, 2018), chickens (Richards *et al.*, 2019) and pigs (Hong *et al.*, 2016). However, some find no effect of phage therapy despite the presence of the host (Zhao *et al.*, 2017). The realm of these discrepancy is not known.

To my knowledge, only two studies have investigated the effect of phage application in environmental ecosystems, finding no changes in the bacterial community in water, but shifts in soil (Pereira *et al.*, 2011; Braga *et al.*, 2020). It is likely that the resolution of the analytical approaches used to study the community (e.g. DGGE vs. Illumina) and the sampling scheme (i.e. temporal resolution and number of sampling sites) influence whether an effect of phages is detected. The published studies suggest that we do not understand when phages induce changes in the community.

No one has investigated how phages affect community assembly patterns and how the density of the target bacterium is related to this. A deeper understanding of the effects of phage treatment is needed.

Bacterial invasions

Invasions are a special case of dispersal and occur when a new population moves into a resident community (Litchman, 2010; Mallon *et al.*, 2015; Kinnunen *et al.*, 2016). Invasions are one of the six major drivers of global biodiversity loss, and it is therefore crucial to understand how invasions affect community assembly (Almond *et al.*, 2022).

Invasions can be either successful or unsuccessful, depending on whether the invader establishes itself in the community (Kinnunen *et al.*, 2016). Successful invasions are thought to affect communities because the invader establishes itself in a niche, thereby affecting bacterial interactions (Mallon *et al.*, 2018). Such successful invasions are of particular concern in areas where humans introduce microorganisms into natural ecosystems (Drake *et al.*, 2007). Even unsuccessful invasions can alter community composition, evenness and richness (Buchberger & Stockenreiter, 2018; Mallon *et al.*, 2018). Therefore, the effects of invasions,

whether successful or unsuccessful, extend beyond the establishment of the invader and may affect community dynamics and diversity.

To my knowledge, no studies have specifically investigated the effect of invasions on community assembly. However, efforts have been made to understand how deterministic and stochastic processes affect the establishment of an invader (Vila *et al.*, 2019). One study found that stochasticity is important for the establishment of an invader (Kinnunen *et al.*, 2018). Furthermore, it has been shown that invasions have an impact on interaction networks (Rivett *et al.*, 2018) and that higher invader **propagule pressures** (i.e. the density of the invader) lead to less similar communities than when propagule pressure is low (Acosta *et al.*, 2015).

It is particularly important to understand how invasions affect communities in situations where bacteria are added to a community in large quantities. For example, to study treatment options to combat bacterial infections or understand probiotic establishment a large number of bacteria are added to an ecosystem (e.g. Borges *et al.*, 2021; Donati *et al*., 2022). When individuals are added to communities where the niches are already occupied, some individuals must die. Nutrients from dead individuals will then temporarily increase the resources available to the community. In this way, invasions could affect community assembly by increasing stochasticity through the same mechanisms as disturbances that increase resource availability. Thus, both successful and unsuccessful invasions can act as disturbances. While specific studies of the effect of invasions on community assembly are lacking, existing research suggests that invasions influence assembly. It is therefore critical to investigate the effects of invasions on community assembly from a nature-conservatory and research-focused perspective.

In summary, Chapter 2 described that any change to the bacterial community or its environment is a disturbance. Understanding how these disturbances affect community diversity, characteristics and assembly patterns is essential to achieve predictable bacterial community management, which is discussed further in Chapter 3.

Chapter 3: Bacterial community management

Chapter 2 introduced that bacterial communities change over time and that we have limited knowledge on how disturbances affect bacterial community composition and assembly patterns. Building on this foundation, Chapter 3 discusses the importance of studying bacterial communities and emphasises their relevance to our society. Using the specific case of land-based aquaculture, this chapter exemplifies how effective management of the bacterial communities in rearing tank water has significantly improved fish viability, highlighting the practical implications and benefits of understanding and managing microbial ecosystems.

Humans and other organisms depend on bacterial communities

Bacterial communities are fundamental to human, animal, and plant life because of our inherent and acquired dependence on bacteria (Figure 10).

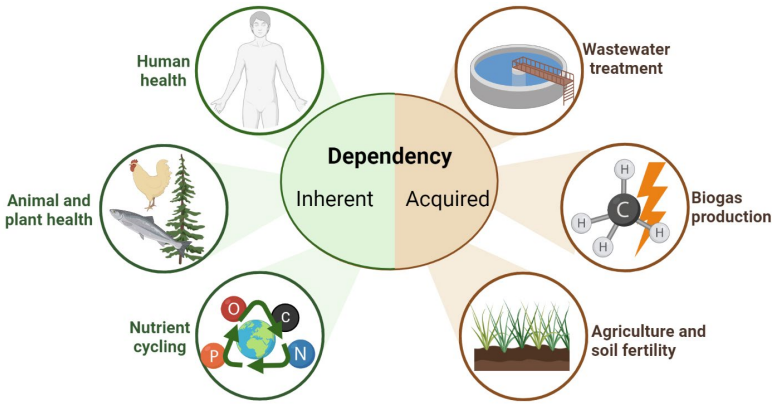


Figure 10: Bacterial communities are essential to many aspects of human life. Bacteria affect the health of humans and other living organisms and are crucial to the biochemical cycling of nutrients on the planet. Furthermore, many industries depend on bacteria, such as the wastewater treatment-, biogas- and agricultural sector.

Inherent dependence on bacterial communities

Humans and other eucaryotes have an inherent dependence on bacterial communities that have shaped the Earth and the organisms that inhabit it since the beginning of life. Bacteria provide numerous ecosystem services that we depend upon (Brandt *et al.*, 2015).

Bacteria are critical to the biogeochemical cycling of nutrients on our planet, affecting our air, land and water. For example, bacteria are critical in the carbon cycle where they break down complex organic compounds into smaller molecules. They are responsible for 25% of the global primary production which fixates carbon dioxide (CO_2) and produces oxygen and are essential for converting nitrogen gas (N_2) to bioavailable ammonia (NH_3) (Kirchman, 2012). All organisms, including humans, have evolved to benefit from the consistent and reliable supply of nutrients provided by bacteria.

Human activities have had such a significant impact on the global ecosystem that our influence is now recognised as a new geological epoch called the Anthropocene (Lewis & Maslin, 2015; Crutzen & Stoermer, 2022). Climate change and contamination have already led to changes in the dynamics of bacterial communities (Doney *et al.*, 2011). For example, CO_2 concentrations in the oceans have increased by about 20%, leading to ocean acidification and reduced marine nitrification (Bates *et al.*, 2012; Hutchins & Fu, 2017). These dramatic human-induced changes in a short period of time are causing numerous disturbances to the bacterial communities on Earth and pose a risk to our mutualistic dependence on bacteria. It is also predicted that there will be a dramatic increase in the occurrence of extreme weather and catastrophic events such as floods, droughts and forest fires, which disrupt bacterial communities (Pörtner *et al.*, 2022). These changes could lead to changes in the global nutrient cycle, potentially altering the availability of nutrients on which organisms depend.

In addition, every known animal and plant is colonised by bacteria and other microorganisms. Bacteria provide beneficial functions to their host, such as producing vitamins, aiding in the digestion and absorption of nutrients, and supporting the immune system and normal development (Zhang *et al.*, 2015). The relationship between bacteria and humans is so intricate that there is a correlation between a gut bacterial community out of balance and human illness (Zhang *et al.*, 2015). Examples of such illnesses include irritable bowel syndrome, allergies and depression (Rachid & Chatila, 2016; Harper *et al.*, 2018; Capuco *et al.*, 2020). It is therefore of concern that modern lifestyles appear to disturb the communities, as the human bacterial community has decreased in diversity and changed in composition compared to the pre-industrialised human (Rosas-Plaza *et al.*, 2022). It has been hypothesised

that these changes make the human microbiota less resilient and stable, potentially leaving modern humans more susceptible to infections (Keesing *et al.*, 2010).

Bacterial communities are, therefore, crucial for the functioning of ecosystems and the well-being of organisms. However, disturbances in bacterial communities pose a risk to our inherent dependence on bacteria, with potential implications for global nutrient availability and host health.

Acquired dependency on bacterial communities

Humans have developed a dependence on functioning bacterial communities on an industrial scale. The projected increase in the world population to 9.7 billion by 2050 emphasises the growing need to address sustainability concerns due to the limited availability of resources (United Nations, 2022). In this context, bacteria can play a critical role in several areas that are essential for achieving the sustainability goals. In particular, their involvement in resource recycling and food production is of great importance (United Nations, 2023).

Water is a fundamental resource, but its availability is becoming limited due to increased water use and contamination. Wastewater treatment can purify and facilitate water reuse from municipal and industrial wastewater (Mannina *et al.*, 2022). Bacteria are crucial in biological wastewater treatment as they metabolise and degrade organic matter and other nutrients (Graham & Smith, 2004). Essential resources such as nitrogen and phosphorus can then be recovered as part of the bacterial biomass, which can be processed to produce high quality fertiliser for the agricultural sector (Mannina *et al.*, 2022). It is estimated that fertiliser production will need to increase 50% by 2050 to feed the world population (Christiaens, 2018). At the same time, the availability of phosphorus is declining, and the chemical fixation of nitrogen is an energy-intensive process (Christiaens, 2018; Alewell *et al.*, 2020). Recovery of these highly valuable resources is therefore essential. Wastewater treatment thus ensures both the reuse of water and the recycling of essential nutrient resources.

Optimal performance of wastewater treatment is achieved under stable operating conditions. However, wastewater often varies in composition and abiotic conditions, leading to repeated disturbances to the bacterial communities. For example, rainfall significantly reduces the efficiency of the treatment process, as rainwater carries dissolved soil particles and

potentially toxic chemicals from above-ground pollution into the sewer pipes (Tram Vo *et al.*, 2014). Thus, rainfall increases loading and water flows, resulting in biomass dilution, shortened hydraulic retention time and increased substrate load in the treatment tanks. These operational changes can shift the composition of the bacterial community from a *K*- to an *r*-selected community, which has been shown to result in a loss of functions such as nitrification, denitrification and phosphorus removal (Vuono *et al.*, 2014; Yin *et al.*, 2022). Conversely, prolonged droughts are expected to reduce substrate loads and water flows. As mentioned above, more extreme weather events are expected due to climate change, so a better understanding of how disturbances affect treatment processes is essential.

Furthermore, humanity is facing an energy crisis for which bacteria can be part of the solution (Thiruselvi *et al*., 2021). For example, food waste can be converted into biogas through the process of anaerobic digestion (Pramanik *et al.*, 2019). In this process, a community of bacteria cooperates to hydrolyse and ferment organic molecules into methane (CH_4) and CO_2. Methane can be used for heating, cooking or converted into electrical energy. In addition, as with wastewater treatment, excess bacterial biomass produced during the process can be processed to produce fertiliser for agriculture (Thiruselvi *et al.*, 2021).

Bacteria also play an important role in agriculture as they are critical for soil fertility and form mutualistic bacteria-plant interactions (Ramakrishna *et al.*, 2019). Therefore, manipulating soil conditions to favour beneficial bacteria is important for increasing crop yields. For example, studies of soil bacterial communities have shown that heavy fertilization and pesticide use reduce the abundance of bacteria that promote plant growth (Bissett *et al.*, 2013). This reduction results in a negative feedback loop where the soil needs more added fertilizer and is prone to harbour bacteria that cause plant disease. Furthermore, management of crop- and soil-microbiota can enhance plant resistance to disease (Sullam & Musa, 2021). Therefore, shifting agricultural practices to favour beneficial plant-bacteria interactions can improve soil quality and crop yields.

These examples highlight some areas where bacteria are key players and can be managed. The contributions of bacterial communities extend beyond our current understanding and are intricately intertwined in areas such as human health, the environment and various

industries. As we continue to explore and understand the complex dynamics of these bacterial communities, it is becoming clear that effective management and conservation of these communities is essential for human well-being and sustainability.

Research-driven microbial management in land-based aquaculture

The research on microbial management in land-based aquaculture systems is an excellent example of the benefits of studying bacterial communities. Fish are intimately exposed to their bacterial environment as they breathe, feed and defecate in the water they live in (De Schryver & Vadstein, 2014). Microbial water quality therefore has a major impact on the health and viability of fish. To enhance fish survival and growth in land-based aquaculture it became a research focus to improve the microbial rearing water quality in the 1990s. Numerous studies over 30 years strongly suggest that different water treatment systems affect the composition of the bacterial communities in the rearing tanks (Vadstein *et al.*, 2018). These observations can be explained by the difference in bacterial density between the incoming water and the rearing tank water, as well as the composition of the bacterial communities before they are introduced to the fish (Figure 11). In this thesis, three different water treatment systems were used to investigate their effects on community characteristics.

The **flow-through system** (FTS) is the most straightforward system. In FTS, the intake water is first treated and disinfected before it is pumped into a header water reservoir and finally to the fish tank (Figure 11). In FTS the rearing water is continuously removed from the rearing tank at the same rate as incoming water is introduced (Vadstein *et al.*, 2018). The bacterial communities in FTS water are immature and experience disturbances. One disturbance is the initial disinfection of the intake water, which considerably reduces the bacterial biomass while the carrying capacity remains the same. Thus, immediately after disinfection, competition for resources is removed and the environment turns *r*-selective. The bacteria that survived disinfection thus experience a surplus of available resources, initiating bacterial growth in the pipes and water reservoir. The next major disturbance occurs when the bacteria move from the water reservoir to the rearing tanks. The carrying capacity in the rearing tank water is higher than in the incoming water due to fish feed, faeces and other organic particles in the rearing tank. As a result, the bacterial community experiences a nutrient pulse situation that initiates *r*-selective bacterial regrowth. This regrowth continues until the bacterial density approaches the carrying capacity (Hess-Erga *et al.*, 2010; Vadstein *et al.*, 2018).

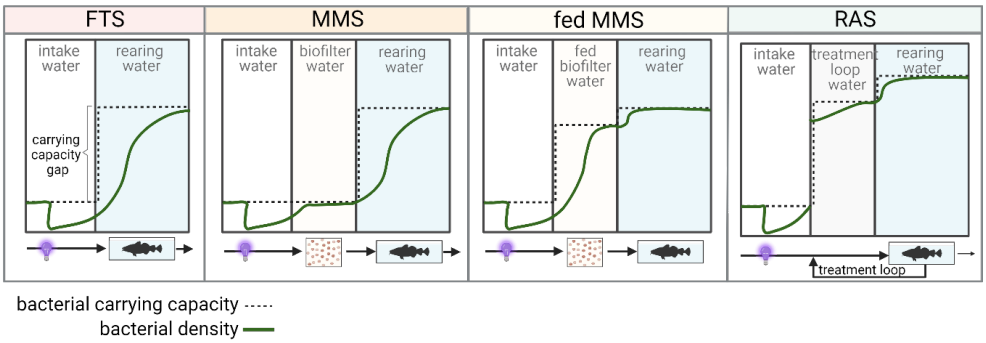


Figure 11: Water treatment configurations in land-based aquaculture systems affect bacterial density and where bacterial regrowth occurs. Initially, UV disinfection of the intake water reduces the bacterial density below the carrying capacity, which initiates bacterial regrowth. In FTS (flow-through system), there is a large carrying capacity gap between the incoming water and the rearing water, and bacterial regrowth takes place in the rearing tank water. In MMS (microbially matured systems), initial growth takes place in a biofilter upstream of the rearing tanks. The bacterial community matures in the biofilm and a mature community is seeded into the rearing tank water. The biofilter can be fed to increase the carrying capacity before entering the rearing tank, thus minimising regrowth in the tank water. In RAS (recirculating aquaculture systems), bacterial regrowth in the rearing tank water is largely eliminated. In RAS, most of the rearing tank water is recirculated and treated in a treatment loop. The carrying capacity in the treatment loop is reduced by the removal of organic material. If there is no UV treatment in the loop, the bacterial density is slightly reduced, and a mature community circulates within the system. Some regrowth takes place when the incoming and treatment loop waters are mixed, but the carrying capacity is reached more quickly, allowing the bacterial community to mature.

The residence time of bacteria in the rearing tank water is relatively low in FTS. Low residence time is expected to result in communities with lower richness and populations with high maximum growth rates (*r*-selection) (Locey & Lennon, 2019). Investigations have shown that the bacterial communities in FTS rearing tank water exhibit growth potential, indicating that carrying capacity has not been reached before the water is exchanged. Thus, there is strong evidence that the rearing tank water is *r*-selective. Indeed, the rearing tank water in FTS contains high proportions of opportunistic *r*-strategic bacteria and is unstable (Attramadal *et al.*, 2014, 2021; Vadstein *et al.*, 2018). These conditions can be detrimental to fish and should be avoided (Perry *et al.*, 2020).

The FTS can be modified to a **microbially matured system** (MMS) by installing a maturing biofilter in the water reservoir upstream of the rearing tanks (Figure 11, Skjermo *et al.,* 1997). In MMS, the bacterial community in the intake water is pre-cultured in a *K*-selective environment with high bacterial competition for resources. The bacterial community in this pre-cultured water thus transitions from a pioneer to a mature community before entering the fish tanks. Thus, although MMS and FTS have a similar gap in carrying capacity between the incoming and rearing water, bacterial regrowth in the rearing tank occurs by a more mature community (Vadstein *et al.*, 2018). The MMS system can further be modified by feeding the biofilter (fed MMS) to close the carrying capacity gap between the intake and rearing water (Figure 11, Attramadal *et al.,* 2016).

These minor modifications to the FTS result in significantly different rearing tank water bacterial communities, which are more stable and contain fewer opportunistic bacteria. For example, Skjermo *et al*. (1997) reported a 55% and Salvesen *et al*. (1999) a 133% reduction in opportunistic bacteria in MMS rearing tank water compared to FTS tank water. In addition, fish survival is significantly higher when reared in mature water than in unmatured water. For example, survival increased when fish were reared in matured instead of unmatured water by 76% for Atlantic halibut (*Hippoglossus hippoglossus*) (Skjermo *et al.*, 1997), 79% and ~100% for turbot (*Scophthalmus maximus*) (Skjermo *et al.*, 1997; Salvesen *et al.*, 1999) and 49% for Atlantic cod (*Gadus morhua*) (Attramadal *et al.*, 2014). These changes strongly suggest that the bacterial community in the rearing tank has a significant impact on fish viability and that MMS facilitates a more beneficial bacterial community for fish (Skjermo *et al.*, 1997; Perry *et al.*, 2020).

The **recirculating aquaculture system** (RAS) is the most complex water treatment system. In RAS, the rearing tank water is recirculated and treated in a treatment loop before it is mixed with treated incoming water (Figure 11, Attramadal *et al*., 2012). Thus, in RAS, the carrying capacity gap between incoming and rearing water is minimal and the degree of bacterial regrowth in the rearing tank water is reduced. In addition, the bacterial communities circulate within the facility which increases the bacterial residence time in the water. Higher residence time is expected to result in communities with greater diversity and higher resource utilisation (*K*-selection) (Locey & Lennon, 2019). Indeed, the rearing tank water in RAS has

29.4% fewer opportunistic bacteria than in FTS suggesting that *K*-selection occurs (Attramadal *et al.*, 2012). Furthermore, the bacterial communities in the rearing tank water are more stable in RAS than in FTS (Attramadal *et al.*, 2014). However, this stability is not achieved if the treatment loop water is disinfected immediately before entering the fish tanks and hydraulic retention time is high (Attramadal *et al.*, 2012, 2021). Studies indicate that the rearing tank water in RAS contains a beneficial bacterial community for the fish, as fish survival is significantly improved in RAS. For example, fish survival in RAS compared to FTS increased by 27% and 52% for Atlantic cod (Attramadal *et al.*, 2012, 2014) and by about 50% for turbot (Verner-Jeffreys *et al.*, 2004). Compared to MMS, RAS has a much higher similarity in bacterial community composition between tanks (Attramadal *et al.*, 2014), indicating that RAS has the highest potential for microbial rearing tank water management of the three systems.

These observations of microbial rearing quality between land-based aquaculture rearing systems suggest the existence of a predictable relationship between operational conditions and bacterial community composition. Consequently, it is crucial to elucidate the processes of bacterial community assembly, which currently represents a significant knowledge gap in larval rearing (Vadstein *et al.*, 2018). By addressing this knowledge gap, it would be possible to develop models that predict how operational changes affect the quality of the bacterial rearing environment. These models can guide management decisions during system disturbances and provide strategies for the recovery of beneficial bacterial communities. Ultimately, such predictive models can help to improve fish health, production outcomes and overall rearing quality.

Managing bacterial communities through predictive models

The ability to manage bacterial communities is highly desirable for human health, ecosystem stability and industrial applications. However, to effectively and appropriately manage these communities, it is crucial to develop accurate models that can predict their response to environmental conditions (De Vrieze *et al.*, 2020). Unfortunately, existing predictive models for the response of bacterial communities to disturbances have low predictability (Zhou & Ning, 2017). This limitation can be attributed to the neglect of stochastic processes during model development (Zhou & Ning, 2017). The incorporation of stochastic processes is

essential to improve the accuracy and utility of these models. Although the outcome of stochastic processes cannot be predicted, it is valuable to understand which environmental conditions that influence stochastic assembly. Such insights allow us to understand how disturbances affect stochastic assembly, so that management practices can avoid these conditions and improve model accuracy (Fukami, 2004; Zhou & Ning, 2017)

Accurate models serve as essential tools for risk assessment. For example, human activities and climate change are predicted to lead to an increase in uncontrollable disturbances (Pörtner *et al.*, 2022). Models can indicate how these disturbances affect bacterial communities and what the ecological consequences might be. As another example, modelling changes in the human gut microbiota following antibiotic treatment can assess the risk of gut dysbiosis (Newton *et al.*, 2023). Conducting appropriate risk assessments to understand the impact of disturbances on our innate and acquired dependencies on bacterial communities is crucial. Accurate models will enable knowledge-based decision-making on bacterial community management strategies and provide guidance for implementing appropriate microbial rescue strategies (Shade, 2023).

In summary, this introduction first described the use of molecular techniques to study bacterial communities, allowing the exploration of their patterns of diversity. These patterns serve as valuable indicators of temporal and disturbance-induced changes in bacterial communities. Secondly, it discussed community assembly processes and what is known about community responses to disturbance. Finally, it shed light on the dependence of humans on bacterial communities, both through our innate and acquired dependency on them. It further highlighted the importance of a comprehensive understanding of community assembly patterns to develop accurate models that can predict community changes and be used to manage bacterial communities. Taken together, these themes set the stage for the investigation of community dynamics explored in this thesis.

Chapter 4: Aims of the thesis

The overall goal of this thesis was to deepen our understanding of bacterial community assembly and to investigate the effects of disturbances on community characteristics such as diversity and composition. The thesis was structured around three sub-goals and included six experiments (Figure 12).

The first goal was to develop a framework to quantify selection and drift in experimental microcosms without dispersal (**Paper I**). This was necessary because currently available analytical tools, such as NTI and βNTI, have underlying assumptions about environmental variability that are often not met in laboratory conditions. In addition, I investigated the potential limitations of the NTI and βNTI frameworks to improve our understanding of their methodological shortcomings (**Work I**).

The second goal was to improve our understanding of bacterial management in rearing tank water in land-based aquaculture. This goal was achieved by investigating the effects of different water treatment systems on bacterial community characteristics in rearing water and larval guts (**Paper II**, **Paper III** and **Paper IV**). Using a disturbance-based framework, these experiments sought to relate community characteristics and assembly processes to the different water treatment systems.

The third goal was to gain a more mechanistic understanding of the effects of disturbances on bacterial community structure and assembly using replicated laboratory microcosms. The disturbances investigated were periodical dilutions resulting in increased resource availability and partial mortality (**Paper I**), selective mortality by antibiotics and bacteriophages (**Paper V** and **Paper VI**), and bacterial invasion (**Paper V** and **Paper VI**).

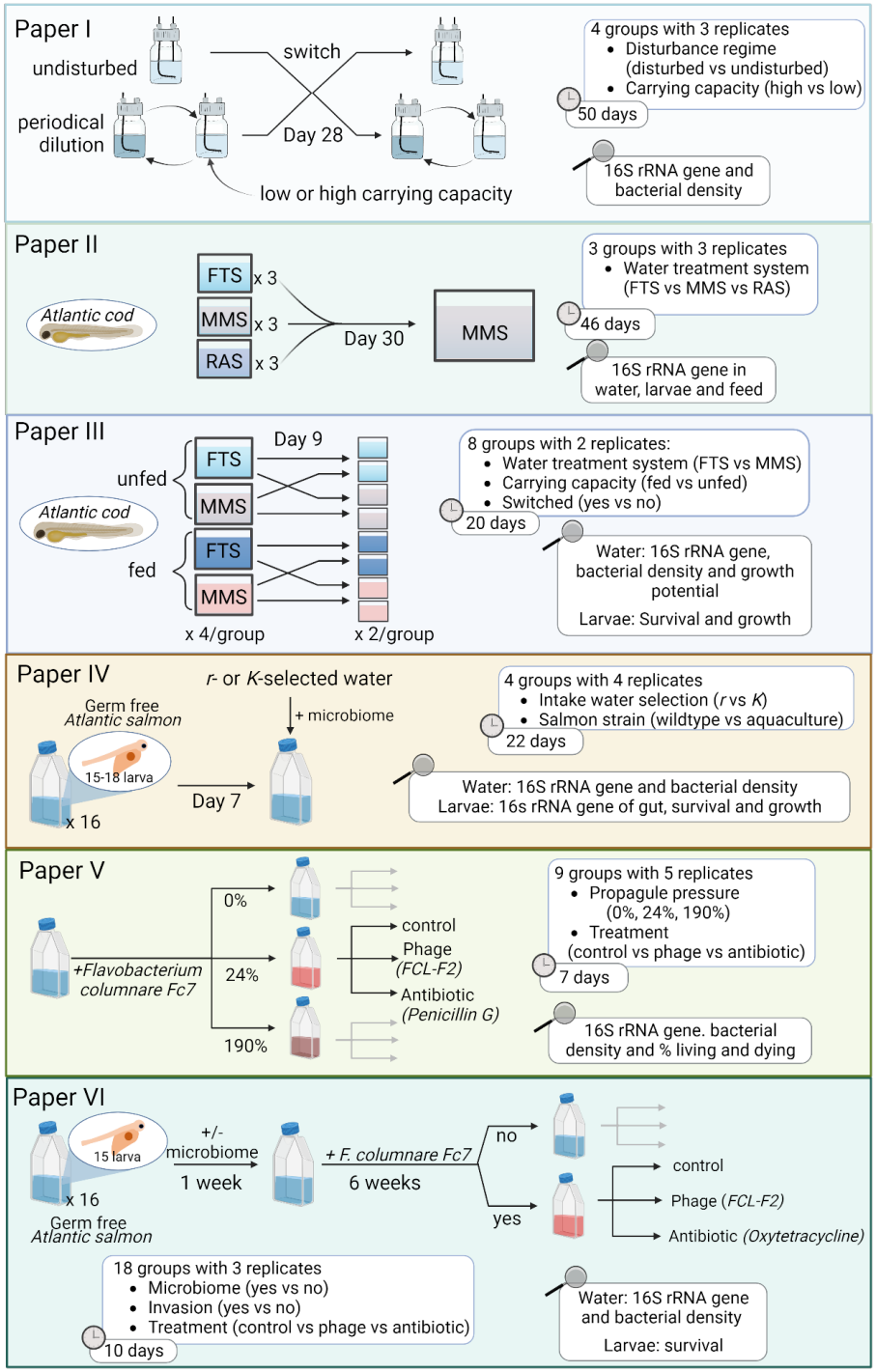


Figure 12: An overview of the experimental design and analytical approaches utilized in Paper I-VI.

Chapter 5: Summary of results and discussion

Quantifying community assembly is challenging

The structure of the phylogenetic tree affects estimates of community assembly

The inherent lack of pattern during stochastic community assembly makes it difficult to assess true ecological stochasticity. Other random effects introduced by the sampling and analytical approach (e.g. amplicon sequencing bias), a lack of insight into relevant deterministic environmental variables, and the overall complexity of bacterial communities make quantifying stochasticity challenging. The currently available frameworks for estimating assembly can be complicated to understand and different frameworks yield different estimates (Sloan *et al.*, 2006; Stegen *et al.*, 2012, 2013; Ovaskainen *et al.*, 2017; Ning *et al.*, 2019, 2020). This inconsistency suggests that the available frameworks are still exploratory rather than conclusive.

Today, analytical approaches such as NTI and βNTI are widely used without consensus or guidelines on how to perform the analysis. What is lacking, in my opinion, is a general guideline on how to define the regional population pool, as the phylogenetic tree is fundamental to the output of the analysis. Therefore, I investigated how the regional population pool affected community assembly in **Work I.**

Specifically, I investigated how errors in the phylogenetic tree and reducing the size of the metacommunity affected the estimates of NTI and βNTI in the datasets from Paper II and Paper IV. In these datasets, the bacterial community was characterised using 16S rRNA gene amplicon sequencing in incoming water, rearing water and fish gut samples. In the initial analysis of the sequencing data from Paper IV, I found that salmon 18S rRNA gene sequences had been classified as bacteria by the bioinformatics pipelines. Such errors are not easily detected by standard microbiota analysis techniques. Therefore, I investigated how such errors affect the null model results. By comparing NTI and βNTI from datasets with and without these outlier sequences, I found that errors in the phylogenetic tree caused significant changes in the conclusions drawn regarding community assembly. These changes were most drastic for the dataset from Paper IV, where 43% of the comparisons of gut samples had so different βNTI estimates that the predicted dominant community assembly

process changed from heterogeneous selection to stochastic. Moreover 66% of the rearing water sample comparisons changed from stochastic to homogeneous selection.

I further reduced the regional population pool in each dataset by omitting sample types. In most publications quantifying NTI or βNTI, the metacommunity is defined as all samples collected. However, we rarely have the opportunity to sample all desired ecosystems due to time or economic constraints. As there is a positive correlation between sample size and richness it is expected that a metacommunity based on several sample types will be more diverse than if only one or a few ecosystems are sampled. I found that reducing the metacommunity size by excluding sample types decreased NTI and increased βNTI. Thus, reducing the metacommunity resulted in estimates closer to the null model. However, compared to errors in the phylogenetic tree, these changes due to metacommunity size were small.

The exploratory analysis presented in **Work I** led to two recommendations. The first and most important recommendation is that the investigator must visualise the phylogenetic tree to confirm that the estimated tree is without outliers. Misclassified eukaryotic and archaeal sequences could affect the average phylogenetic distances in the tree and significantly affect the phylogenetic analysis. Even if the assembly is not examined, visualisation of the phylogenetic tree can be valuable as an aid to data quality assurance. A pipeline to remove outlier sequences in R is added as a complement to this thesis. The introduction of erroneous sequences can be minimised by adding expected eucaryotic or archaeal rRNA gene sequences to the reference database prior to taxonomic classification.

The second recommendation is that NTI and βNTI close to the boundary between deterministic and stochastic should be examined as metacommunity size affects the estimates. Manipulating the metacommunity size and reanalysing samples close to -2 and 2 in NTI or βNTI should be done to confirm whether the conclusions are robust or due to the metacommunity size.

'Replicate similarity rate of change' as a framework for quantifying selection and drift in controlled experimental microcosms

Investigating community assembly in well-replicated ecosystems has been identified as essential for understanding community assembly dynamics (Zhou & Ning, 2017). Null models, such as NTI and βNTI, have been developed for natural ecosystems with gradients in physiochemical environmental conditions. Therefore, the underlying assumptions of the null models are not appropriate in replicated laboratory microcosms, where environmental variation is minimal by design. Furthermore, the inability to account for drift from other stochastic processes is a major drawback of null model approaches. Thus, in **Paper I,** we developed a new analytical framework to estimate drift and selection in replicated ecosystems without dispersal (Figure 13).

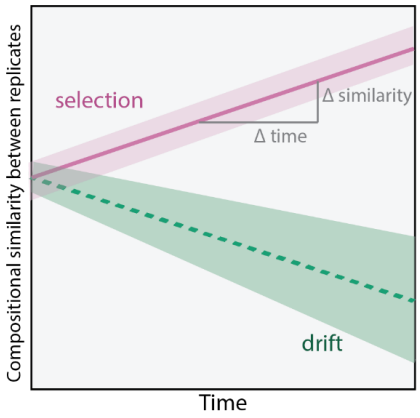


Figure 13: The 'replicate similarity rate of change' framework is based on quantifying community similarity between replicate microcosms at each sampling time and estimating the slope of the temporal change in similarity. If selection structures community assembly, communities should become more similar over time and the variation in community similarity between replicate microcosms should decrease. In contrast, if drift structures community assembly, community composition should diverge over time and variation in community composition should increase.

The proposed framework is based on quantifying the rate of change in similarity between replicate systems without dispersal. The underlying premise of the framework is that all biological replicates originate from the same community and have the same environmental conditions. According to first principal community assembly theory, if selection dominates community assembly, pairs of replicates should increase in similarity over time. Conversely, if drift dominates community assembly, pairs of replicate communities should decrease in similarity over time. Thus, the rate of change is the slope of a regression that explains replicate similarity as a function of time. Further, the variation in replicate similarity at each timepoint can further be assessed to strengthen the conclusions based on the rate of change. It is

expected that variation between replicates should increase over time if assembly is structured by drift, and conversely be stable or decrease if selection dominates.

The 'replicate similarity rate of change' framework has several advantages. First, the framework is more intuitive than null model approaches because it is based on the fundamental principals of ecological assembly theory. Secondly, the framework makes no assumptions about a phylogenetic signal, which can be difficult to estimate accurately for bacterial communities (Münkemüller *et al.*, 2012). Thirdly, the similarity index used to quantify change between replicate communities can easily be changed depending on which community properties one wishes to emphasise. For example, Sørensen similarity is suitable for assessing the degree of competitive exclusion due to drift, whereas Bray-Curtis better reflects fluctuations in abundance.

However, there are some critical limitations to the framework. The most important limitation is that it does not assess dispersal or diversification. Therefore, the framework should only be used in replicated biological systems with complete dispersal limitation and on a timescale where diversification can be expected to have a minimal impact. The bioinformatics approach is expected to influence the contribution of diversification. For example, OTU clustering at 97% will be less influenced by diversification than at the ASV level. Therefore, studies over long timescales should consider agglomerating the population table to a higher taxonomic rank to ensure that the assumption of negligible diversification is maintained. It should be noted that agglomeration of data will affect the confidence of estimates of community assembly processes, as information will be lost.

Furthermore, as with null model approaches, replicate similarity is a community-wide assembly analysis. Recently it has been suggested, and it is reasonable to assume, that different populations follow individual community assembly patterns. New analytical tools have been developed that bin populations in a community based on phylogenetic properties before performing null-based model analysis (e.g. iCAMP, Ning *et al.,* 2020). Such a population-specific focus can be implemented in the 'replicate similarity rate of change' framework. This implementation can be achieved by subsetting the data for specific populations that are thought to be similarly structured prior to performing the analysis.

It is expected that new analytical tools to infer ecological community assembly will be developed, as no method can statistically disentangle diversification, drift, weak selection and weak dispersal. Gaining a comprehensive understanding of all assembly processes is of great value, particularly in understanding how environmental conditions affect bacterial communities. This knowledge goes beyond fundamental research and has practical applications in various industrial settings, such as land-based aquaculture, where the management of the bacterial communities in the rearing tank water is crucial. By applying the knowledge gained from assembly analysis, land-based aquaculture operations can optimise their practices and improve the overall rearing environment for fish larvae.

The bacterial quality of rearing tank water needs continuous management

To obtain a *K*-selected rearing environment, the incoming water and rearing water must have a similar carrying capacity

Previous studies have demonstrated that the water treatment system in land-based aquaculture influences the bacterial communities in the rearing tank water (Vadstein *et al.*, 2018). For instance, when rearing tank water contains an excess of available resources for bacteria, it promotes *r*-selection, which favours bacteria with high maximum growth rates. These conditions promote bacterial regrowth and can lead to detrimental interactions between bacteria and fish. Instead, a better microbial rearing tank water quality is obtained with *K*-selection. *K*-selection occurs when bacterial densities approach or reach the carrying capacity of the ecosystem. In such situations, bacteria must compete for resources, which over time selects for more competitive and specialised *K*-strategist bacterial populations. A likely reason for why *K*-selection improves rearing quality is that competition reduces the likelihood of proliferation of opportunistic bacteria that can be harmful to fish. However, there is still a knowledge gap regarding how *K*-selection is achieved in rearing waters in practice (Vadstein *et al.*, 2018). The results from **Paper II**, **Paper III** and **Paper IV** demonstrate the importance of keeping a similar carrying capacity between the rearing tank water and incoming water to obtain *K*-selection within the rearing environment.

In **Paper II,** Atlantic cod larvae were reared in FTS, MMS, or RAS for 30 days before all tanks received MMS water until day 46 (Figure 12). We observed that the bacterial community

composition of the rearing tank water was more similar between the FTS and MMS than the RAS (Figure 14a). Quantification of bacterial density and bacterial production indicated substantial bacterial growth within the rearing tanks in the FTS and MMS. These findings indicated that both FTS and MMS promoted *r*-selection (data presented in Attramadal *et al.,* 2014). The density ratio between the rearing and incoming water was, on average, 4.2 in the FTS, 6.8 in the MMS and 1.8 in the RAS (Figure 14d). In the RAS there was a small difference in bacterial density between incoming and rearing water and low internal bacterial production which indicated a more *K*-selective environment in the rearing tank water.

These observations suggested that small differences in bacterial density between the incoming and rearing water is determining for obtaining *K*-selection in the rearing tank water. We confirmed this suggestion in **Paper III** where we reared Atlantic cod in fed or unfed FTS and MMS. In the fed systems, the carrying capacity was increased by adding organic material directly to the FTS rearing tanks or to the MMS biofilter. After nine days, the water treatment system in half of the tanks was changed by switching the inlet pipes (Figure 12). We concluded that only the fed MMS promoted *K*-selection by quantifying bacterial density and growth potential and characterising the bacterial community composition.

Firstly, the bacterial communities in the rearing tank water in the unfed FTS and MMS showed growth potential and clustered together in the PCoA ordinations which indicated that these tanks were under a similar selection regime (Figure 14c). No growth potential was measured in the fed systems, indicating that carrying capacity had been reached. Secondly, there were substantial differences in bacterial densities between incoming and rearing water in the different systems (Figure 14d). For example, on the second day, the density ratio between rearing and incoming water was only 1.2 in the fed MMS, 2.16 in the unfed MMS, 2.08 in the unfed FTS and a staggering 132 in the fed FTS. Thus, most of the bacterial growth occurred in the biofilter for the fed MMS and in the rearing tank water for the other systems.

The difference in bacterial regrowth in the rearing tank water between the fed FTS and MMS is particularly intriguing. These systems had a similar carrying capacity in the rearing tank

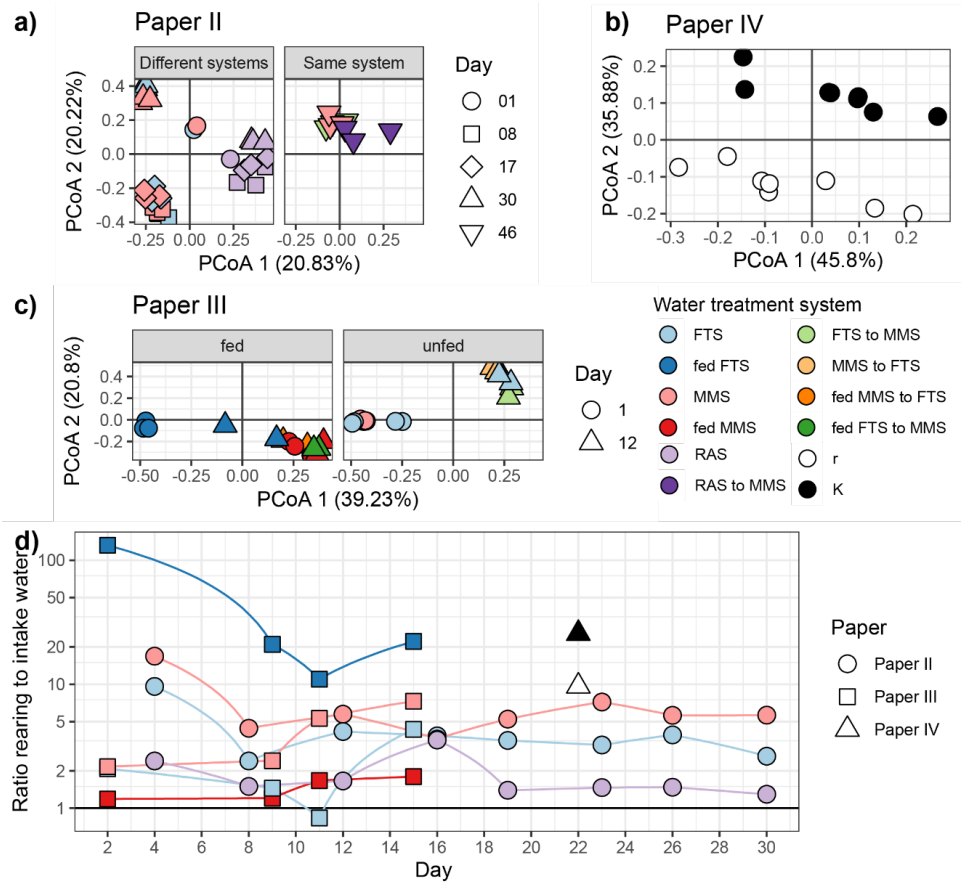


Figure 14: Bray-Curtis based PCoA ordinations of the rearing water in the experiments from a) Paper II, b) Paper IV and c) Paper III. d) The average ratio difference in bacterial density over time between the rearing and intake water in Paper II-Paper IV.

water. However, the bacterial communities in the fed FTS had a substantially lower bacterial density in the incoming water than the rearing tank water and experienced a surplus of resources when moving into the rearing tank. Thus, this finding clearly shows that *r*- and *K*-selection is facilitated by the experienced available resources and not by the carrying capacity itself. These results are in agreement with Attramadal *et al*. (2016), who investigated the differences in bacterial communities between fed and unfed MMS and concluded that only fed MMS facilitated *K*-selection in the rearing tank water (Attramadal *et al.*, 2016).

These conclusions were further supported in **Paper IV,** where Atlantic salmon (*Salmo salar)* were reared in an *r*- or *K*-selected lake-derived bacterial community in a gnotobiotic model system (Figure 12). The bacterial communities in the *r*- and *K*-selected water differed significantly before being added to the fish flasks. Compared to the *r*-selected water, the *K*-selected water had 1.3x higher richness and 1.2x higher evenness. Higher richness and evenness are characteristics of more mature systems (Vadstein *et al.*, 2018). However, when *K*-selected water was added to the fish flask, the composition and diversity changed drastically. The Bray-Curtis similarity between the bacterial communities in the added water and rearing flask water was only 0.06, and evenness decreased 1.6x. The bacterial density ratio was 25.6 in the *K*-selected and 9.6 in the *r*-selected flask water, indicating that carrying capacity was higher in the rearing flask water and that all flask water was *r*-selected (Figure 14d).

These results add to our understanding of the relationship between carrying capacity and bacterial regrowth in the rearing water. They show that higher carrying capacities, and hence bacterial densities, can be beneficial for obtaining *K*-selected rearing water, but only if bacterial regrowth in the rearing water is minimised.

Different rearing environments create different stable states

The observations in Paper II, Paper III and Paper IV provide additional evidence that the water treatment system impacts the bacterial communities in the rearing tank water (Vadstein *et al.*, 2018). We found that the different systems create different stable states for the rearing water bacterial communities (Figure 14a,c). The Bray-Curtis based PCoA ordinations of these communities indicate that *K*-selected rearing water reach a stable state more rapidly. For instance, the RAS and fed MMS had a relatively stable community compositions over time and achieved this composition after just one day (Figure 14a,c). This stability suggests that a mature community is seeded from the biofilter to the rearing tank, which facilitates colonisation of a *K*-selected community in the rearing tank water. Moreover, switching the water treatment system changed the bacterial community composition in the rearing tank water. For example, in **Paper II**, all systems were changed to the MMS after 30 days, which resulted in increased similarity in community composition between the tanks (Figure 14a).

Similarly, in **Paper III**, the bacterial community composition shifted to be more similar to the stable state in the water treatment system they changed to (Figure 14c).

Three important conclusions can be drawn from these results. Firstly, operational conditions such as the water treatment system and available resources affect the bacterial communities in the rearing tank water. This emphasises the importance of managing these conditions to promote a favourable microbial rearing environment for the fish. Secondly, for systems without water recirculation, the time required to achieve a stable bacterial community in the rearing tank water can be accelerated by increasing the carrying capacity of the incoming water, for example by adding organic matter to biofilters (i.e. fed MMS). This means that targeted adjustments of the rearing environment have the potential to establish a beneficial bacterial community more quickly in the rearing tank water, which impacts fish health and production outcomes. Finally, changes in rearing water conditions affect the composition of the bacterial community in the rearing tank water. Therefore, rearing water management is a continuous process.

Community assembly patterns in rearing water and fish are correlated

The different water treatment systems affected the degree of stochastic and deterministic assembly processes in the bacterial communities of the rearing tank water and larval guts. The communities in the rearing tank water was phylogenetically structured in all samples from Paper II and Paper IV (NTI>2, Figure 15a). This clustering indicated that niche filtering occurred in the rearing tank water.

In **Paper II,** we quantified the degree of stochastic and deterministic processes in the rearing tank water within each water treatment system over time (Figure 15b). We observed the greatest differences in βNTI between systems at the beginning of the experiment, where RAS had the highest degree of stochasticity, followed by FTS and then MMS. The average βNTI decreased in all systems until day 17, indicating a greater contribution of deterministic processes in assembly of the bacterial communities in the rearing tank water. Thereafter, βNTI increased, and there was no difference in βNTI between systems. This observation indicated a higher contribution of stochastic processes at days 30 and 46 (Figure 15b).

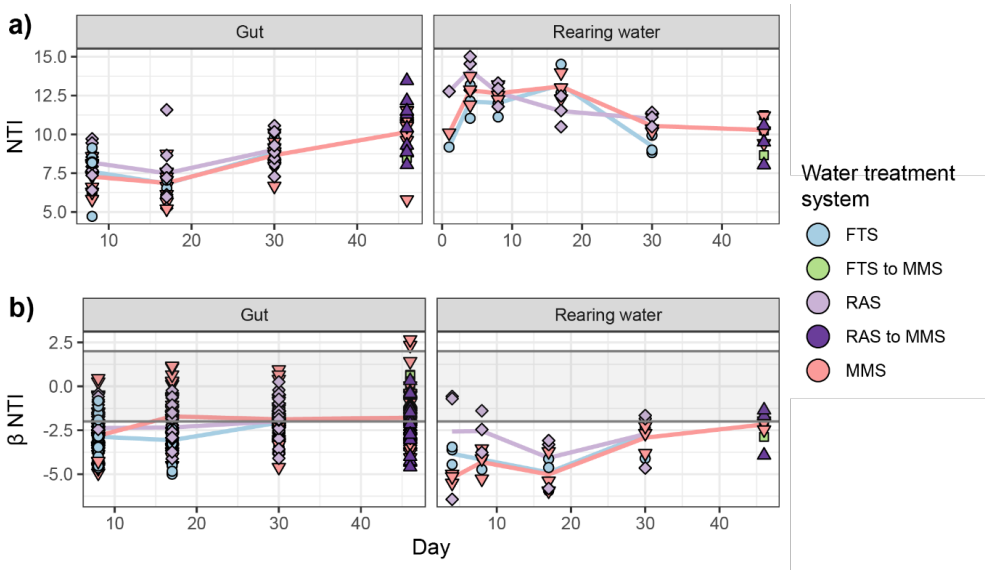


Figure 15: The a) NTI and b) βNTI in the fish and water samples in Paper II on the different sampling days. Each point represents a sample (a) or sample comparison (b) within a water treatment system. The line indicates the average NTI or βNTI. The grey shaded area indicates where community assembly is stochastic as the values are not significantly different from the null model.

Furthermore, we found that although the gut bacterial communities in individual larvae were phylogenetically clustered, the community composition varied considerably between individual cod larvae and salmon alevins. Such inter-individual variation has previously been observed in Atlantic cod and suggested to be due to stochastic processes (Fjellheim *et al.*, 2012; Bakke *et al.*, 2015). Indeed, we found that stochastic processes contributed to inter-individual community variation, as many βNTI were not significantly different from the null model (Figure 15b). In **Paper II**, the percentage of cod gut comparisons classified as stochastic increased over time and varied between 22-64%. There were significant differences in the assembly of the gut bacterial communities between the rearing systems only at day 17 (ANOVA $p<0.05$), otherwise the assembly pattern was similar. In **Paper IV,** we found that at day 20, 42% and 69% of the salmon gut bacterial communities were structured by stochastic processes when reared in *K*-selected and *r*-selected incoming water, respectively. Thus, these studies contribute to the growing evidence that stochastic processes are important in the assembly of bacterial communities (Zhou & Ning, 2017).

In another study of cod at 0, 8 and 12 weeks of age, the contribution of stochastic processes increased with development (Keating *et al.*, 2022). Further, a study of salmon at the parr, smolt and adult life stages found that stochastic processes almost completely dominated the assembly of the gut bacterial communities (Heys *et al.*, 2020). Community assembly was characterised throughout the life cycle of gibel carp (*Carassius auratus gibelio*), in which deterministic processes decreased substantially over developmental stages (Li *et al.*, 2017). Taken together, these observations of the assembly of gut bacterial communities at larval, juvenile and adult life stages suggest a shift from deterministic to stochastic assembly during development. Exactly when this shift occurs is unknown. For salmon, we do not have a detailed understanding of the initial community assembly processes during the alevin stages. While community assembly in cod has been characterised in more detail during the larval stages (weeks 1, 2, 4 and 7 in Paper II and 0, 8 and 12 in Keating *et al*., 2022), we know little about community assembly in the cod gut during the juvenile and adult life stages. Future studies should aim to understand these patterns throughout the life cycle to determine whether stochasticity increases continuously as fish develop into adulthood.

A fascinating finding was that while the bacterial community composition in the rearing tank water and cod gut samples in **Paper II** was significantly different, there was a positive correlation between the assembly processes in the rearing tank water and gut samples (Pearson r = 0.38, p = 0.0005, Figure 16). Although the variation in gut βNTI was large, there is a trend that as the stochasticity in the rearing tank water community increases, the stochasticity in the gut communities also increase. However, we cannot determine whether the fish change the community assembly in the water or whether the water affects the assembly in the fish. Given that stochasticity increases with fish development, it can be speculated that there initially is a strong niche filter within the cod larvae guts, which diminishes over time as bacteria from the feed and water disperse into the fish and compete for colonisation. Since dispersal from the rearing environment to the larvae is a stochastic process, this competition will reduce the compositional similarity of the larvae over time and create more stochastic patterns.

To validate the correlation, an experiment could be designed in which rearing tanks are stocked with different densities of fish, such as low, medium and high densities. In tanks with

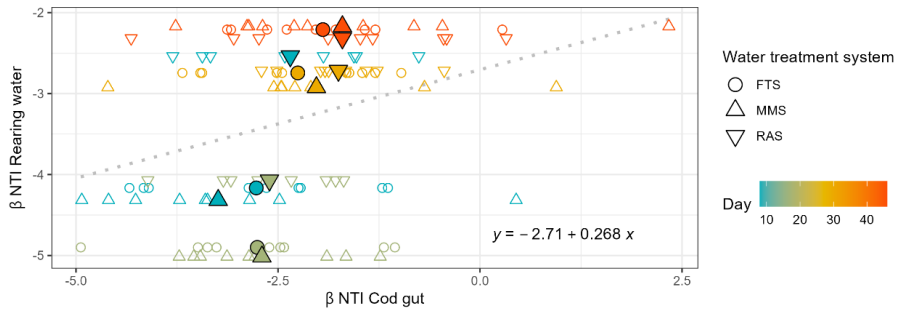


Figure 16: The βNTI for cod gut sample comparisons from the same rearing tank versus the βNTI in the rearing tank water. The colours indicate the sampling day. Note that each βNTI from a rearing tank water sample is compared to several larvae gut βNTI as 2-3 larvae were sampled from each tank. The dotted line indicates the linear regression with equation y = -2.71+0.268x. The filled points are the average within each rearing system at each sampling timepoint. The linear regression of the average points has an equation of y = 0.88+1.82x and thus is much steeper.

higher stocking densities, the increased number of fish should result in more frequent dispersal events between the fish and water. If the observed correlation in assembly is valid, tanks with high fish densities should exhibit a stronger correlation in assembly patterns. To gain insight into the cause-effect relationships underlying this correlation, an experiment could be conducted in which the history of the rearing tank water is manipulated. For example, fish could receive rearing tank water either directly from the water treatment system (representing standard operation) or from a system containing older fish. Based on our and previous findings, the rearing tank water from the tanks with older fish should have higher stochasticity. If the assembly patterns of the bacterial community in the rearing tank water have an effect on the assembly in the bacterial communities in the fish, it is expected that stochasticity is higher in the fish receiving water from tanks containing older fish than those reared under standard operating conditions. As a control, a group without fish could be included to assess community assembly in the rearing water alone. Such experiments can assess the consistency and causality of the observed correlation and shed light on the interplay between community assembly processes in fish and water.

In summary, the fish larvae and the rearing tank water have phylogenetically clustered bacterial communities, indicating that environmental filtering is important for the structure of the bacterial communities. However, stochastic processes also contribute to community

assembly. The observations indicate that the contribution of stochastic assembly in bacterial communities of the rearing tank water and fish guts increases with time. This increase is possibly due to a reduction of niche filtering within the gut over time, combined with stochastic dispersal of bacteria from the rearing water to the fish.

***K*-selected water benefits larval viability, and management is a continuous process**

The different water treatment systems affected the survival and viability of cod larvae in Paper II and Paper III. In **Paper II**, the survival of cod was recorded 30 days after hatching and showed that the RAS and MMS had 65% higher survival than the FTS (Attramadal *et al.*, 2014). Because the physicochemical water quality was similar between the FTS and MMS it was speculated that the bacterial rearing quality impacted the larval viability. The water going into the RAS and MMS was *K*-selected, while only RAS remained *K*-selected in the rearing tanks. Due to differences in carrying capacity between the incoming and rearing water, the bacterial community in the MMS and FTS rearing tank water was under *r*-selection. When the MMS incoming water is transferred to the rearing tank water with higher carrying capacity bacterial regrowth and selection for *r*-strategists occurs. However, as the bacterial community in the incoming MMS water was matured and *K*-selected in the biofilter, the abundance of the *r*-strategist populations was likely low in the incoming water. The difference in bacterial density between the incoming and rearing tank water in the FTS and MMS allowed for approximately 2-3 cell divisions before carrying capacity was reached (Attramadal *et al.*, 2014). Thus, in the MMS, the regrowth of bacteria in the *r*-selective rearing tank water environment likely resulted in a community with fewer opportunistic bacteria than the FTS due to a more favourable and *K*-selected bacterial community in the incoming water. Thus, the composition of the bacterial community in the incoming water appears to have an impact on larval viability.

Attramadal *et al*. (2014) further hypothesised that the larval survival was higher in MMS and RAS than FTS because these systems facilitated for a more beneficial initial colonisation of the larvae. Therefore, in **Paper III,** we specifically investigated whether the initial rearing environment leave legacy effects in the bacterial community in the rearing tank water or in the cod.

The main conclusion of **Paper III** was that no legacy effects were found in neither the cod larvae nor the bacterial community of the rearing water. Instead, a consistent finding was that if a rearing tank was connected to the fed FTS at any time, survival was extremely low, even down to 0%. In contrast, survival in the other systems was much higher, ranging from 13% to 23%. The difference in carrying capacity between the incoming and rearing tank water was greatest in the fed FTS, suggesting that *r*-selection was stronger in these tanks than in MMS and FTS. This observation suggests that strong *r*-selection induces more detrimental fish-bacteria relationships than weak *r*-selection. Furthermore, while cod larvae reared continuously in the fed MMS had an average survival of 16%, survival in tanks that switched from fed FTS to MMS was only 1%. These findings on survival disprove that initial conditions leave protective legacies in the larvae (Attramadal *et al.*, 2014). Instead, continuous microbial management of the rearing water is critical to secure good larval viability.

This lack of a legacy effect was also observed in Nile tilapia (*Oreochromis niloticus*), where differences in the gut bacterial community caused by early rearing in FTS or MMS (bio-floc system) diminished when larvae were transferred to a RAS (Deng *et al.*, 2021). Similarly, the bacterial community in the zebrafish gut does not show historical effects of initial rearing conditions. Instead, the developmental stage of the fish appears to be the primary determinant of community composition (Xiao *et al.*, 2021). The importance of developmental stage has also been demonstrated for grass carp (*Ctenopharyngodon idellus*), mandarin fish (*Siniperca chuatsi*) and Chinese largemouth catfish (*Silurus meridionalis*) (Yan *et al.*, 2016). Thus, across a range of fish species, it appears that the niches in the fish change during development, resulting in a continuously changing environment to which the bacteria must adapt.

The Atlantic cod gut must be colonised with bacteria from the rearing environment. However, little is known about the dispersal of bacteria from the rearing environment to the fish and their subsequent colonisation and establishment within the gut. For example, in zebrafish, gut bacterial communities have been found to become more dispersal-limited with increasing age, suggesting that dispersal between zebrafish and water and among zebrafish decreases over time (Burns *et al*., 2015). This decrease may indicate that host selection of bacteria decreases with age, which is consistent with the absence of legacy effects in fish and increased

stochastic assembly of the gut communities. As the knowledge on how bacteria from the rearing environment disperse and establish in the fish is poor more research is needed. Future research should aim to elucidate how initial gut colonisation and subsequent succession within the gut occur. From an applied perspective, such knowledge will be highly influential during the initial larval stages where most of the mortality occurs.

In summary, Paper II, Paper III and Paper IV investigated the relationship between the water treatment system and the characteristics of the bacterial communities in the rearing tank water and fish guts as well as fish viability. We found that *K*-selected rearing tank water can be achieved by ensuring that the carrying capacity between the incoming and rearing tank water is similar. Our results show that the water treatment systems and available resources in the water changes the bacterial community in the rearing tank water. Therefore the microbial rearing quality needs to be continuously managed. Furthermore, stochastic community assembly increases over time in bacterial communities of the rearing water and cod gut. Whilst deterministic assembly is still a major contributor to assembly, this increase in stochasticity may indicate that microbial rearing water management needs to be monitored more closely as the fish develop.

It was evident that the water treatment system influenced the bacterial communities in the rearing tank water. From the perspective of the bacterial community, the environmental conditions in FTS and MMS are more unstable than in RAS, as resource availability and biomass concentrations change throughout the system pipelines. Therefore, the bacterial community experiences more disturbance in the FTS and MMS, and we subsequently observed that the degree of disturbance affected the bacterial community. To better understand the effect of disturbance on community assembly, we induced disturbance under controlled laboratory conditions.

Periodical dilution increased the contribution of selection

Studying bacterial community assembly in replicated ecosystems with identical environmental conditions has been emphasised as essential to gain a detailed understanding of the assembly dynamics (Zhou & Ning, 2017). In **Paper I,** the effect of periodical dilutions

and carrying capacity on community assembly was investigated using a replicated microcosm setup. The experiment had a 2x2 factorial crossover design which varied the disturbance regime (disturbed/undisturbed) and carrying capacity (high/low) and had a switch in the disturbance regime halfway through the experiment (Figure 12). The main aim of **Paper I** was twofold. Firstly, we wanted to elucidate how periodical dilutions impacted the contribution of selection and drift, and whether the disturbance history of the community influenced these trends. Secondly, previous studies from land-based aquaculture found that the carrying capacity in the rearing tank water impacted bacterial community composition (e.g. Attramadal *et al*., 2016, Paper II and Paper III). However, to my knowledge, how carrying capacity impact community assembly had not been investigated.

Using the 'replicate similarity rate of change' framework, we found that undisturbed bacterial communities were predominantly structured by drift, while disturbed communities were structured by selection. Carrying capacity had no significant effect on community assembly or composition. This lack of effect may be due to the relatively small 5x difference in carrying capacity. Recently, carrying capacity was shown to have a large effect on the stability of a synthetic bacterial community consisting of 48 different populations subjected to periodic dilutions in media with low, medium or high nutrient concentrations (Hu *et al.*, 2022). In the low nutrient concentration, the communities were stable and replicate similarity was high. However, stability was not achieved in the medium or high nutrient concentrations, which had approximately 10x and 40x higher carrying capacities than the low condition (Hu *et al.*, 2022). I therefore speculate that we would have observed a greater contribution of drift in the high carrying capacity group if the difference in carrying capacity was of a larger magnitude. Future experiments should investigate the effect of carrying capacity on community characteristics, as nutrient availability is an essential variable for bacterial growth and the absolute number of bacteria the environment can sustain. Identifying the boundary where resource availability facilitates stability and where it does not is of great value.

The switch in disturbance regime was valuable because it allowed us to speculate on the stable states of community composition. We identified two locally stable attractors in community composition associated with the two disturbance regimes. Remarkably, after the switch in disturbance regime, the community compositions exhibited a regime shift in the

stable state, changing from the undisturbed stable state to the disturbed state, or vice versa. Without the switch in disturbance regime, this directional and reversible effect of disturbance on bacterial community composition would not have been observed. These observations are similar to the shifts in community composition from the stable state when the water treatment system was changed in Paper II and Paper IV. Thus, our results strongly suggest that the environmental conditions in the water influences the community composition and this makes microbial management possible.

We found that the periodical dilutions increased the contribution of selection and shifted community composition to more *r*-strategic bacteria such as *Vibrio* and *Colwellia* (Liu *et al.*, 2017). We speculate that stochastic effects were suppressed by the periodical dilutions that resulted in exponential growth periods due to density-independent biomass loss combined with resource input. These results corroborate the observations in Mao *et al*. (2023), where a planktonic bacterial community was split into replicate microcosms and pulse disturbed with various degrees of dilutions. Higher dilution factors enhanced deterministic processes and increased the average number of rrn gene copies populations in the communities had. Here we showed that periodical dilutions (i.e. repeated pulse disturbances) also induce these community assembly patterns.

Studies of soil bacterial communities also show that disturbances that cause partial mortality increase the average number of rrn gene copies of the community (Nemergut *et al.*, 2015; Kearns & Shade, 2018). Furthermore, nutrient enrichment as a disturbance shifts community assembly from deterministic to stochastic (Yang *et al.*, 2018), and carbon enrichment decreases dissolved oxygen concentrations, indicating *r*-strategist blooms (Garnier *et al.*, 2017). Dilutions can be considered as a dual disturbance because both mortality and the concentration of available resources increase. We observed that the periodical dilutions increased the contribution of the deterministic process selection. Because disturbances that cause mortality shifts assembly to be more deterministic, while nutrient enrichment increases stochasticity, it appears that mortality is a stronger predictor in structuring community assembly than nutrient enrichment is. To test this hypothesis one could investigate how assembly patterns change following dilutions with solvents with or without available

resources. Such an experiment can disentangle the dual disturbance effect of dilutions and enhance our understanding of which underlying mechanisms that affect assembly.

In summary, we found that periodic dilutions increased the contribution of selection, whereas bacterial communities with a stable environment were more structured by drift. We speculate that increased partial mortality, combined with excess resources following dilution, selected for *r*-strategic communities. Next, we investigated how disturbances by unsuccessful invasion affected assembly.

Unsuccessful invasions increased the contribution of drift

Although several papers have investigated the effect disturbances have on increasing the probability of successful invader establishment (e.g. Liu *et al.,* 2012; Lear *et al.,* 2022), investigating bacterial invasions as ecosystem disturbances is relatively underexplored. To my knowledge, only Amor *et al*. (2020) have analysed invasions using a disturbance framework. They found that unsuccessful invasions can profoundly affect community composition, leading to a shift towards a new alternative stable state. While previous studies have shown that increased propagule pressure of the invader significantly impact community composition (Acosta *et al.*, 2015; Albright *et al.*, 2020), these studies did not use a disturbance framework. Therefore, in **Paper IV** and **Paper V,** we aimed to explore the role of invasion as a disturbance, and in **Paper IV** we assessed the effects of varying the propagule pressure of the invader.

In **Paper IV** and **Paper V**, we introduced the gram-negative bacterium *Flavobacterium columnare* strain Fc7 into freshwater bacterial communities to study the consequences of invasions (Figure 12). Despite propagule pressures of 24%, 190% and 500%, the invasions were ultimately unsuccessful. By the end of the studies (7 days Paper IV and 10 days Paper V), the abundance of the invader had fallen to around 0%, meaning that *F. columnare* had not established. Not surprisingly, the community composition in both experiments shifted one day after the invasion event due to the substantial addition of the *F. columnare* strain.

Interestingly, in **Paper V**, we observed significant differences in community characteristics between the invaded and non-invaded communities one week after invasion. These

differences indicated that the unsuccessful invasions had disturbed the community beyond the duration of the invasion event. For example, the invaded communities had 1.3x higher richness (linear model, $p<0.05$, adj. $R^2 = 0.81$), significantly different community compositions (pairwise PERMANOVA $p<0.05$, $R^2>0.70$) and community composition variation (PERMDISP $p<0.001$). In **Paper VI**, however, such invasion-induced changes were not observed as successional patterns in α-diversity and community composition were similar between uninvaded and invaded communities.

Previous studies have focused on understanding the role of stochastic and deterministic processes in the establishment of invaders (Kinnunen *et al.*, 2018). However, the impact of invasions on bacterial community assembly has received limited attention and remains largely unexplored. **Paper V** sheds light on this aspect, showing that unsuccessful invasions cause a shift in community assembly from selection to drift, with a more pronounced shift observed under higher propagule pressures of the invader.

We postulate that this shift in assembly was driven by ecological factors related to nutrient dynamics. Specifically, we observed that *F. columnare* densities decreased more rapidly than if growth was zero, indicating that cells were dying in the microcosms. We speculate that F. *columnare* lysed and released resources that the resident community could use as substrate for growth. In addition, the initial lake-derived inoculum was filtered at 55 µm which allowed small protozoa to be included in the microcosms. Although we did not confirm protozoa presence, they could have played a role in facilitating bacterial turnover through grazing, potentially contributing to the recycling of resources from *F. columnare* (Kirchman, 2012).

Although we did not measure dissolved organic matter concentrations the potential release of resources from dead *F. columnare* cells appears likely as we observed a feast-famine response in bacterial density. The feast-famine response occurs when a community exploits a sudden burst of nutrients leading to an increase in cell density. However, due to resource depletion, the increased densities cannot be sustained, and followingly the density declines rapidly (Himeoka & Mitarai, 2020). Moreover, we observed that community assembly shifted from selection to drift following the invasion. Similar shifts from deterministic to stochastic assembly have been observed with nutrient pulses (e.g. Santillan & Wuertz, 2022; Zhou *et al.,*

2014). I therefore propose that unsuccessful invasions where the invader dies act as a disturbance in a similar way as nutrient pulses do.

It is important to note that both **Paper IV** and **Paper V** were conducted as semi-continuous systems. In semi-continuous systems, pulses of increased resources remain available to the community until they are consumed or the inoculum is diluted. Thus, the potential release of resources from *F. columnare* cells can be exploited by the community. In contrast, in continuous systems, medium is constantly added to the inoculum at the same rate as excess volume is removed. The flow rate describes how fast the volume within a continuous system is removed. Thus, in a continuous system and depending on the flow rate, released *F. columnare* resources might be removed from the community before they can be exploited. Therefore, the observed community changes due to the potential release of dead *F. columnare* cell resources may have been very different in a continuous system. On the other hand, in a batch system where there is no input or removal of inoculum, the effects might have been greater. Comparing the effects of invasions in batch, semi-continuous and fully continuous systems would be an exciting avenue for future research.

The propagule pressures of the invader used in these experiments were significantly higher than those expected to occur naturally. Thus, future studies should investigate the effects of invasions at realistic propagule pressures to improve our understanding of the ecological dynamics associated with bacterial invasions under realistic invasion scenarios.

Successful invasions result in the establishment of the invader, which can lead to regime shifts or changes in the stable state of ecosystems (Acosta *et al.*, 2015). Here, I have shown that unsuccessful invasions, although not resulting in establishment, still have ecological consequences, as dead cells are likely to lyse and introduce a nutrient pulse disturbance. We observed that in semi-continuous systems, the unsuccessful invasion event led to an increased contribution of drift. In the next section, it will be discussed that antibiotics have many of the same disturbance effects as unsuccessful invasions.

Selective mortality by a bactericidal antibiotic increased drift

Antibiotics are known to disturb bacterial communities, generally resulting in altered community composition and reduced diversity compared to unexposed communities (Grenni *et al.*, 2018). During exposure, antibiotics inhibit growth and potentially kill populations susceptible to the chemical. Thus, one would initially expect determinism to increase due to targeted mortality and increased selection pressure for resistant bacteria (Pallecchi *et al.*, 2008). However, recently, it has been reported that drift increases at the community level following antibiotic exposure (Chen *et al.*, 2023). As a matter of fact, not much attention has been paid to how antibiotics affect community assembly. Therefore, in **Paper IV** and **Paper V**, antibiotics were added as a disturbance to understand their effects on community characteristics and assembly in planktonic bacterial communities.

In **Paper IV**, we observed the drastic effects antibiotics can have on community characteristics. In this experiment, Penicillin G was added as a one-time treatment. At the community level, addition of antibiotics significantly reduced richness and evenness and induced significant changes in the composition of the bacterial community compared to the control. For example, after one week, richness was almost halved in the antibiotic-treated microcosms compared to the control. There was also evidence for that antibiotic exposure was a strong enough disturbance to drive the communities into an alternative stable state. These changes in community characteristics are similar to those observed in a previous study (Antonopoulos *et al.*, 2009), and suggest that Penicillin G disturbs planktonic bacterial communities.

We also observed that antibiotics increased the contribution of drift compared to the control treatment. This shift was most pronounced in terms of population abundance (Bray-Curtis based), but was also evident in terms of population presence (Sørensen based). This observation is consistent with the fact that some antibiotic-susceptible individuals survive exposure to antibiotics and thereafter exhibit stochastic growth (Coates *et al.*, 2018). As bacterial density, richness and proportion of living cells decreased after antibiotic exposure, I speculate that the community assembly patterns were affected by similar mechanisms to those we observed for different invasion propagule pressures. That is, bacteria died and

subsequently released resources to the rest of the community, resulting in a nutrient pulse. However, in contrast to the death of only one invader (i.e. *F. columare*), the Penicillin G antibiotic treatment caused the death of more bacterial types. In addition to potentially increasing resource turnover, the death of established populations affects the complex network of community interactions. This means that the competitive and mutualistic relationships of which the dying bacteria are part of cease. These immediate effects lead to a relaxation of niche competition within the community (Chen *et al.*, 2023). Such relaxation can increase the contribution of drift by facilitating the growth of previously suppressed populations. Thus, while antibiotics have a deterministic effect in reducing the abundance of antibiotic-susceptible populations, the effect on community assembly may manifest stochastically.

Paper V and Chen *et al.* (2023) suggest that antibiotic treatment increases the contribution of drift in dispersal-limited systems. However, not all studies find this pattern. For example, in the bacterial communities in the gut of spider mites and in lettuce soil, no changes were detected in assembly due to antibiotic exposure (Shen *et al.*, 2021; Zhu *et al.*, 2023). Thus, the assembly shifted in aquatic ecosystems but remained unchanged in more structured environments such as gut and soil. It is possible that changes in community assembly depend on the type of ecosystem. Quantification of community assembly in more ecosystem types is needed to fully understand when antibiotics increase the contribution of drift.

The ISH proposes that intermediate levels of disturbance should increase stochastic community assembly, as the advantage of having specialised traits diminishes in such unpredictable environments (Santillan *et al.*, 2019). This increased stochasticity is hypothesised to be reflected in α-diversity, which is predicted to reach its maximum at higher stochastic contributions. We observed the opposite in **Paper V**, where α-diversity decreased as drift increased. It is important to note that our study included only a single dose of antibiotics, which limits the direct applicability of our results to the validation of ISH. Nevertheless, our observations highlight the need for experiments using different antibiotic concentrations and realistic treatment frequencies, to investigate whether responses to antibiotic exposure are consistent with the predictions of the ISH.

We observed a significant increase in *γ-proteobacteria* after antibiotic exposure. In **Paper V**, this increase appeared to be at the expense of *β-proteobacteria*, which decreased significantly. In particular, we observed a substantial bloom of the genus *Pseudomonas* known to contains many pathogenic strains (Peix *et al.*, 2009). Many *γ-proteobacteria* are classified as *r*-strategists (Fuhrman & Hagström, 2008). Thus, the changes in community composition support that the antibiotics caused an increase in resource availability, as antibiotics appeared to induce *r*-selection.

That Penicillin G created an *r*-selective environment is of clinical concern. An already compromised host in need of medical treatment such as antibiotics does not benefit from an *r*-selective environment, which increases the likelihood of secondary infections. Thus, our results add to the growing accumulation of evidence that antibiotics can have negative consequences for animals and humans (Patangia *et al.*, 2022).

Penicillin G is a bactericidal antibiotic that causes cell lysis. Bactericidal antibiotics have been shown to inhibit bacterial growth differently than bacteriostatic antibiotics (Coates *et al.*, 2018). Therefore, in **Paper VI**, we used the bacteriostatic antibiotic oxytetracycline, which reversibly inhibits protein synthesis. Oxytetracycline is widely used in the aquaculture sector (Lulijwa *et al.*, 2020), and has been shown to affect the gut microbiota of juvenile salmon (Navarrete *et al.*, 2008).

In **Paper VI,** oxytetracycline was added daily. Curiously, oxytetracycline did not cause a significant decrease in bacterial density, reduction in α-diversity or changes in bacterial community composition compared to the control treatment. Thus, oxytetracycline did not appear to disrupt the bacterial community, which was unexpected. Because oxytetracycline causes growth inhibition and several water exchanges were performed during the experiment we expected the abundance of the *F. columnare* strain to decline. Although we confirmed that the *F. columnare* Fc7 strain was susceptible to oxytetracycline prior to the experiment, we did not observe a reduction in the relative 16S rRNA gene copies of the invader. Quantifying the percentage of living and dying populations would have been valuable in this experiment to observe absolute changes in the community dynamics. The lack of a disturbing effect of oxytetracycline on bacterial community characteristics has also been reported in soil

(Unger *et al.*, 2013). We speculate that bacterial turnover was so low that the effect of growth inhibition was irrelevant in this experiment and therefore had little effect on the community.

Since oxytetracycline appeared to be ineffective, I will only conclude on the effects of Penicillin G as a community disturbance. The results suggest that drift increases when antibiotics are introduced into systems without continuous water exchange, most likely due to increased available resources to antibiotic resistant populations. Future experiments could test this hypothesis by culturing one or more antibiotic-sensitive populations with labelled carbon, nitrogen and phosphorus and test whether these elements are transferred to other populations following antibiotic treatment. In addition, quantifying the concentration of dissolved organic matter would be a valuable parameter to correlate with bacterial density measurements to confirm whether nutrients become available to the rest of the community. Finally, it has been observed that antibiotics increase the contribution of deterministic processes when added to continuously cultured communities (Eckert *et al.*, 2023). Thus, it may be that when antibiotics induce mortality in batch or semi-continuous systems, the effect is like a disturbance in resource availability which enhances stochastic processes. In continuous systems, resources from dead cells might be removed from the ecosystem before the community can exploit the nutrient increase. Therefore, deterministic processes may be enhanced as all antibiotic-susceptible populations are removed from the community, thus freeing niches and initiating strong competition between the remaining populations. Such a scenario is similar as described in Zhou *et al*. (2014) where disturbances that enhance mortality are hypothesised to increase the contribution of deterministic processes. A deeper understanding of how community assembly is affected in batch, semi-continuous and continuous systems would be an exciting avenue for future research.

In summary, the antibiotic Penicillin G increased the contribution of drift, significantly altered bacterial density, richness and composition, and created an *r*-selective environment. I speculate that these changes were driven by increased resource availability caused by selective mortality of antibiotic-susceptible populations. Oxytetracycline did not cause any observable changes in community characteristics such as diversity and composition as compared to the control. The effects of selective mortality on community characteristics and

composition are, however, not limited to antibiotic treatment. This brings us to the next topic of investigation: the effects of population-selective mortality with phages.

Selective mortality with phages does not affect community assembly

Phages have been proposed as a suitable alternative to antibiotics due to their high selectivity towards specific bacterial types and bactericidal properties. However, the effects of phages on bacterial community properties and assembly patterns are poorly understood. We shed light on these effects in **Paper V** and **Paper VI**. In these papers, we added the phage FCL-2 to bacterial communities with and without the phage host *F. columnare Fc7.*

When the host *F. columnare* was absent from the community, we found no effects of phage treatment on bacterial community characteristics in **Paper VI**. However, minor effects were observed in **Paper V.** For example, after one week the phage-treated bacterial communities had higher densities and a shift in community composition as compared to the control. This shift was attributed to a significant increase in the genus *Aquirufa*. Members of this genus is known to be specialists in remineralising organic matter (Reichenbach, 2006). The increased bacterial density and growth of *Aquirufa* suggests that some community members were able to grow using the added phages as resources. Although studies are scarce, there is evidence for that heterotrophic bacteria can utilise phage particles as a substrate for growth (Noble & Fuhrman, 1999; Noble *et al.*, 1999). Thus, future experiments should investigate how viral decay might induce changes to bacterial community properties.

When the bacterial host *F. columnare* was present in **Paper V** or **Paper VI**, we did not observe any significant changes in bacterial community characteristics or assembly patterns compared to the control treatment. This lack of an effect was a surprising finding as we had speculated that lysis of the bacterial host *F. columnare* would increase nutrient availability to the resident community. However, a drawback of Paper V and Paper VI was that *F. columnare* failed to establish in the community. As a result, *F. columnare* abundances also decreased in the untreated controls. As discussed previously, increased propagule pressure of *F. columnare* resulted in a shift from selection to drift. Unfortunately, in Paper V we were unable to disentangle the effect of phage treatment from the effect of unsuccessful invasions.

Therefore, although we could not statistically prove that phages increased the contribution of drift, it is highly likely that phages would significantly increase the contribution of drift if the bacterial host were to proliferate in the control group.

Phages did not affect the bacterial community characteristics compared to the control treatment. For example, we found no evidence for that phage treatment selected for *r*-strategic bacteria, as we observed with antibiotic treatment. Therefore, phage treatment is a better alternative to antibiotics for protecting the bacterial community structure.

Phages can be a suitable prophylactic tool to keep pathogenic bacteria out of ecosystems. For example, in Paper II, we observed that an OTU classified as *Arcobacter* had low abundance in the rearing tank water but high abundance in the bacterial community of the cod gut. Members of *Arcobacter* are known to be pathogenic to fish (Fitzgerald & Nachamkin, 2015). Therefore, this observation suggests that fish pathogens may flourish when they enter their host, despite being present at low abundances in the fish's environment. It is important to keep possible pathogenic strains out of the aquatic microbiota to reduce the risk of fish infection. In **Paper VI,** Atlantic salmon alevins with and without a microbiome were reared in water invaded or not with *F. columnare* (Figure 12). We observed that larval survival was higher when phages were added to the rearing water invaded with *F. columnare* than in the invaded control treatment. Thus, this work adds to the growing body of evidence that phages do not harm eukaryotic organisms (Oliveira *et al.*, 2012). Similar observations that phage treatment increase survival have been observed in shrimp (Karunasagar *et al.*, 2007), Atlantic salmon (Higuera *et al.*, 2013), artemia (Kalatzis *et al.*, 2016), and a myriad of other fish species (Oliveira *et al.*, 2012). Furthermore, the phage FCL-2 has been shown to survive in RAS facilities for up to 21 days, even in the absence of the host *F. columnare* (Almeida *et al.*, 2019). Thus, phages may be a suitable tool to achieve pathogen-free water and increase fish production.

While Paper V and Paper VI have increased our understanding of the effect of phage treatments on community characteristics, there are still many knowledge gaps that need to be addressed. Firstly, as suggested for further comprehension of the effect of invasions and antibiotics, investigations of the effect of phage treatments on community characteristics

should be carried out and compared between batch, semi-continuous and continuous systems. Secondly, these studies were conducted in communities where the phage host*, F. columnare*, was not an established community member. As such, it is doubtful that it was part of the complex interaction networks that bacterial communities have. Removing a population that is a critical node in such networks could potentially have drastic cascading effects on community characteristics. Such cascading effects have been demonstrated in 10-species assemblages (Hsu *et al.*, 2019) and should be further explored in natural ecosystems and more complex communities.

Overall, we found that phage treatment significantly altered community characteristics when the phage host *F. columnare* was absent. When the host was present, we observed no changes in community characteristics compared to the control. The bacterial host *F. columnare* failed to establish in the communities and was eliminated in both the phage treatment and control groups. It is therefore likely that changes can be detected if established hosts are targeted by the phage. Nevertheless, compared to antibiotics, phage treatment caused significantly less disruption to the bacterial communities. Therefore, phage treatment is proving to be a suitable alternative to antibiotics.

Conclusion

The goal of this thesis was to investigate bacterial community characteristics and assembly in relation to disturbances. I first demonstrated that the structure of the phylogenetic tree influences the interpretation of the NTI and βNTI and recommend that the tree is examined to ensure that it includes only bacteria. In addition, some assumptions underlying NTI and βNTI are not met under controlled laboratory conditions. To address this, we successfully developed the 'replicate similarity rate of change' framework, which quantifies selection and drift in dispersal-limited replicate microcosms. This framework was used to investigate the effects of periodic dilution, unsuccessful invasion, Penicillin G antibiotic treatment and phage treatment on community assembly in semi-continuous microcosms. The results showed that periodic dilutions increased the contribution of selection, while unsuccessful invasions and antibiotic treatment increased the contribution of drift. These disturbances led to a shift in community composition towards an alternative stable state, likely driven by increased resource availability that favoured bacterial growth and *r*-selection. Notably, phage treatment did not alter community assembly patterns or characteristics.

Furthermore, the effects of different water treatment systems on the bacterial communities in the rearing tank water and fish larvae were investigated. We found that the water treatment systems played a crucial role in the assembly of the rearing water bacterial community and that maintaining a similar carrying capacity between the incoming and rearing water was critical for promoting beneficial *K*-selection. Interestingly, stochastic processes increased over time in both fish larvae and the rearing water, suggesting that microbial management of the rearing water may require more detailed attention as the fish develop.

In summary, this work provides important insights into the characteristics and assembly of bacterial communities in response to ecosystem disturbances. It identifies several key knowledge gaps that require further investigation to gain a mechanistic understanding of the changes in community assembly. Addressing these gaps is essential for the development of accurate models that can effectively capture the dynamics of bacterial communities and their responses to environmental change.

References

Acosta, F., Zamor, R. M., Najar, F. Z., Roe, B. A., & Hambright, K. D. (2015). Dynamics of an experimental microbial invasion. *PNAS*, *112*(37), 11594–11599. doi: 10.1073/pnas.1505204112

Ahasan, Md. S., Kinobe, R., Elliott, L., Owens, L., Scott, J., Picard, J., Huerlimann, R., & Ariel, E. (2019). Bacteriophage versus antibiotic therapy on gut bacterial communities of juvenile green turtle, Chelonia mydas. *Environmental Microbiology*, *21*(8), 2871–2885. doi: 10.1111/1462-2920.14644

Albright, M. B. N., Sevanto, S., Gallegos-Graves, L. V., & Dunbar, J. (2020). Biotic Interactions Are More Important than Propagule Pressure in Microbial Community Invasions. *MBio*, *11*(5). doi: 10.1128/mBio.02089-20

Alewell, C., Ringeval, B., Ballabio, C., Robinson, D. A., Panagos, P., & Borrelli, P. (2020). Global phosphorus shortage will be aggravated by soil erosion. *Nature Communications*, *11*(1), 4546. doi: 10.1038/s41467-020-18326-7

Almeida, G. M. F., Mäkelä, K., Laanto, E., Pulkkinen, J., Vielma, J., & Sundberg, L.-R. (2019). The Fate of Bacteriophages in Recirculating Aquaculture Systems (RAS)—Towards Developing Phage Therapy for RAS. *Antibiotics*, *8*(4), 192. doi: 10.3390/antibiotics8040192

Almond, R. E. A., Grooten, M., Juffe Bignoli, D., & Petersen, T. (2022). *WWF: Living Planet Report 2022 – Building a nature-positive society*.

Amor, D. R., Ratzke, C., & Gore, J. (2020). Transient invaders can induce shifts between alternative stable states of microbial communities. *Science Advances*, *6*(8). doi: 10.1126/sciadv.aay8676

Andrews, J. H., & Harris, R. F. (1986). r- and K-Selection and Microbial Ecology. In *Advances in Microbial Ecology* (Vol. 9, pp. 99–147). doi: 10.1007/978-1-4757-0611-6_3

Antonopoulos, D. A., Huse, S. M., Morrison, H. G., Schmidt, T. M., Sogin, M. L., & Young, V. B. (2009). Reproducible Community Dynamics of the Gastrointestinal Microbiota following Antibiotic Perturbation. *Infection and Immunity*, *77*(6), 2367–2375. doi: 10.1128/IAI.01520-08

Attramadal, K. J. K., Minniti, G., Øie, G., Kjørsvik, E., Østensen, M.-A., Bakke, I., & Vadstein, O. (2016). Microbial maturation of intake water at different carrying capacities affects microbial control in rearing tanks for marine fish larvae. *Aquaculture*, *457*, 68–72. doi: 10.1016/j.aquaculture.2016.02.015

Attramadal, K. J. K., Øie, G., Størseth, T. R., Alver, M. O., Vadstein, O., & Olsen, Y. (2012). The effects of moderate ozonation or high intensity UV-irradiation on the microbial environment in RAS for marine larvae. *Aquaculture*, *330–333*, 121–129. doi: 10.1016/j.aquaculture.2011.11.042

Attramadal, K. J. K., Øien, J. V., Kristensen, E., Evjemo, J. O., Kjørsvik, E., Vadstein, O., & Bakke, I. (2021). UV treatment in RAS influences the rearing water microbiota and reduces the survival of European lobster larvae (Homarus gammarus). *Aquacultural Engineering*, *94*, 102176. doi: 10.1016/j.aquaeng.2021.102176

Attramadal, K. J. K., Salvesen, I., Xue, R., Øie, G., Størseth, T. R., Vadstein, O., & Olsen, Y. (2012). Recirculation as a possible microbial control strategy in the production of marine larvae. *Aquacultural Engineering*, *46*, 27–39. doi: 10.1016/j.aquaeng.2011.10.003

Attramadal, K. J. K., Truong, T. M. H., Bakke, I., Skjermo, J., Olsen, Y., & Vadstein, O. (2014). RAS and microbial maturation as tools for K-selection of microbial communities improve survival in cod larvae. *Aquaculture*, *432*, 483–490. doi: 10.1016/j.aquaculture.2014.05.052

Baas Becking, L. G. M. (1934). *Geobiologie of inleiding tot de milieukunde*. W.P. Van Stockum & Zoon.

Bakke, I., Coward, E., Andersen, T., & Vadstein, O. (2015). Selection in the host structures the microbiota associated with developing cod larvae (Gadus morhua). *Environmental Microbiology*, *17*(10), 3914–3924. doi: 10.1111/1462-2920.12888

Bar-On, Y. M., Phillips, R., & Milo, R. (2018). The biomass distribution on Earth. *PNAS*, *115*(25), 6506–6511. doi: 10.1073/pnas.1711842115

Bates, N. R., Best, M. H. P., Neely, K., Garley, R., Dickson, A. G., & Johnson, R. J. (2012). Detecting anthropogenic carbon dioxide uptake and ocean acidification in the North Atlantic Ocean. *Biogeosciences*, *9*(7), 2509–2522. doi: 10.5194/bg-9-2509-2012

Bent, S. M., Miller, C. A., Sharp, K. H., Hansel, C. M., & Apprill, A. (2021). Differential Patterns of Microbiota Recovery in Symbiotic and Aposymbiotic Corals following Antibiotic Disturbance. *MSystems*, *6*(2). doi: 10.1128/mSystems.01086-20

Berg, G., Rybakova, D., Fischer, D., Cernava, T., Vergès, M.-C. C., Charles, T., Chen, X., Cocolin, L., Eversole, K., Corral, G. H., Kazou, M., Kinkel, L., Lange, L., Lima, N., Loy, A., Macklin, J. A., Maguin, E., Mauchline, T., McClure, R., … Schloter, M. (2020). Microbiome definition re-visited: old concepts and new challenges. *Microbiome*, *8*(1), 103. doi: 10.1186/s40168-020-00875-0

Bissett, A., Brown, M. V., Siciliano, S. D., & Thrall, P. H. (2013). Microbial community responses to anthropogenically induced environmental change: towards a systems approach. *Ecology Letters*, *16*(SUPPL.1), 128–139. doi: 10.1111/ele.12109

Borges, N., Keller-Costa, T., Sanches-Fernandes, G. M. M., Louvado, A., Gomes, N. C. M., & Costa, R. (2021). Bacteriome Structure, Function, and Probiotics in Fish Larviculture: The Good, the Bad, and the Gaps. *Annual Review of Animal Biosciences*, *9*(1), 423–452. doi: 10.1146/annurev-animal-062920-113114

Boughner, L. A., & Singh, P. (2016). Microbial Ecology: Where are we now? *Postdoc Journal*, *4*(11), 3. doi: 10.14304/SURYA.JPR.V4N11.2

Braga, L. P. P., Spor, A., Kot, W., Breuil, M.-C., Hansen, L. H., Setubal, J. C., & Philippot, L. (2020). Impact of phages on soil bacterial communities and nitrogen availability under different assembly scenarios. *Microbiome*, *8*(1), 52. doi: 10.1186/s40168-020-00822-z

Brandt, K. K., Amézquita, A., Backhaus, T., Boxall, A., Coors, A., Heberer, T., Lawrence, J. R., Lazorchak, J., Schönfeld, J., Snape, J. R., Zhu, Y.-G., & Topp, E. (2015). Ecotoxicological assessment of antibiotics: A call for improved consideration of microorganisms. *Environment International*, *85*, 189–205. doi: 10.1016/j.envint.2015.09.013

Buchberger, F., & Stockenreiter, M. (2018). Unsuccessful invaders structure a natural freshwater phytoplankton community. *Ecosphere*, *9*(3). doi: 10.1002/ecs2.2158

Burns, A. R., Stephens, W. Z., Stagaman, K., Wong, S., Rawls, J. F., Guillemin, K., & Bohannan, B. J. (2016). Contribution of neutral processes to the assembly of gut microbial communities in the zebrafish over host development. *The ISME Journal*, *10*(3), 655–664. doi: 10.1038/ismej.2015.142

Bush, K., & Bradford, P. A. (2016). β-Lactams and β-Lactamase Inhibitors: An Overview. *Cold Spring Harbor Perspectives in Medicine*, *6*(8), a025247. doi: 10.1101/cshperspect.a025247

Cairns, J., Jokela, R., Becks, L., Mustonen, V., & Hiltunen, T. (2020). Repeatable ecological dynamics govern the response of experimental communities to antibiotic pulse perturbation. *Nature Ecology & Evolution*, *4*(10), 1385–1394. doi: 10.1038/s41559-020-1272-9

Callahan, B. J., McMurdie, P. J., & Holmes, S. P. (2017). Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *The ISME Journal*, *11*(12), 2639–2643. doi: 10.1038/ismej.2017.119

Capuco, A., Urits, I., Hasoon, J., Chun, R., Gerald, B., Wang, J. K., Kassem, H., Ngo, A. L., Abd-Elsayed, A., Simopoulos, T., Kaye, A. D., & Viswanath, O. (2020). Current Perspectives on Gut Microbiome Dysbiosis and Depression. *Advances in Therapy*, *37*(4), 1328–1346. doi: 10.1007/s12325-020-01272-7

Chanishvili, N. (2012). Phage Therapy—History from Twort and d'Herelle Through Soviet Experience to Current Approaches. In *Advances in virus research* (Vol. 83, pp. 3–40). Adv Virus Res. doi: 10.1016/B978-0-12-394438-2.00001-3

Chase, J. M. (2007). Drought mediates the importance of stochastic community assembly. *PNAS*, *104*(44), 17430–17434. doi: 10.1073/pnas.0704350104

Chase, J. M., & Myers, J. A. (2011). Disentangling the importance of ecological niches from stochastic processes across scales. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*(1576), 2351–2363. doi: 10.1098/rstb.2011.0063

Chen, Z., Zhang, W., Peng, A., Shen, Y., Jin, X., Stedtfeld, R. D., Boyd, S. A., Teppen, B. J., Tiedje, J. M., Gu, C., Zhu, D., Luo, Y., & Li, H. (2023). Bacterial community assembly and antibiotic resistance genes in soils exposed to antibiotics at environmentally relevant concentrations. *Environmental Microbiology*, *25*(8), 1439–1450. doi: 10.1111/1462-2920.16371

Chopra, I., & Roberts, M. (2001). Tetracycline Antibiotics: Mode of Action, Applications, Molecular Biology, and Epidemiology of Bacterial Resistance. *Microbiology and Molecular Biology Reviews*, *65*(2), 232–260. doi: 10.1128/mmbr.65.2.232-260.2001

Christiaens, M. (2018). Technologies for resource recovery from human urine : terrestrial and space applications [Ghent University]. In *Universiteit Gent*. http://hdl.handle.net/1854/LU-8582887

Clark, J. S., Dietze, M., Chakraborty, S., Agarwal, P. K., Ibanez, I., LaDeau, S., & Wolosin, M. (2007). Resolving the biodiversity paradox. *Ecology Letters*, *10*(8), 647–659. doi: 10.1111/j.1461-0248.2007.01041.x

Coates, J., Park, B. R., Le, D., Şimşek, E., Chaudhry, W., & Kim, M. (2018). Antibiotic-induced population fluctuations and stochastic clearance of bacteria. *ELife*, *7*. doi: 10.7554/eLife.32976

Coma, R., Ribes, M., Serrano, E., Jiménez, E., Salat, J., & Pascual, J. (2009). Global warming-enhanced stratification and mass mortality events in the Mediterranean. *PNAS*, *106*(15), 6176–6181. doi: 10.1073/pnas.0805801106

Connell, J. H. (1978). Diversity in Tropical Rain Forests and Coral Reefs. *Science*, *199*(4335), 1302–1310. doi: 10.1126/science.199.4335.1302

Crutzen, P. J., & Stoermer, E. F. (2000). The 'Anthropocene'. *IGBP Global Change Newsletter*, *41*, 17–18. https://www.degruyter.com/document/doi/10.12987/9780300188479-041/html

De Schryver, P., & Vadstein, O. (2014). Ecological theory as a foundation to control pathogenic invasion in aquaculture. *The ISME Journal*, *8*(12), 2360–2368. doi: 10.1038/ismej.2014.84

de Vries, F. T., & Shade, A. (2013). Controls on soil microbial community stability under climate change. *Frontiers in Microbiology*, *4*(SEP), 265. doi: 10.3389/fmicb.2013.00265

De Vrieze, J., De Mulder, T., Matassa, S., Zhou, J., Angenent, L. T., Boon, N., & Verstraete, W. (2020). Stochasticity in microbiology: managing unpredictability to reach the Sustainable Development Goals. *Microbial Biotechnology*, *13*(4), 829–843. doi: 10.1111/1751-7915.13575

Deng, Y., Kokou, F., Eding, E. H., & Verdegem, M. C. J. (2021). Impact of early-life rearing history on gut microbiome succession and performance of Nile tilapia. *Animal Microbiome*, *3*(1), 81. doi: 10.1186/s42523-021-00145-w

Dini-Andreote, F., Stegen, J. C., van Elsas, J. D., & Salles, J. F. (2015). Disentangling mechanisms that mediate the balance between stochastic and deterministic processes in microbial succession. *PNAS*, *112*(11), E1326 LP-E1332. doi: 10.1073/pnas.1414261112

Dion, M. B., Oechslin, F., & Moineau, S. (2020). Phage diversity, genomics and phylogeny. *Nature Reviews Microbiology*, *18*(3), 125–138. doi: 10.1038/s41579-019-0311-5

Donati, V. L., Madsen, L., Middelboe, M., Strube, M. L., & Dalsgaard, I. (2022). The Gut Microbiota of Healthy and Flavobacterium psychrophilum-Infected Rainbow Trout Fry Is Shaped by Antibiotics and Phage Therapies. *Frontiers in Microbiology*, *13*, 987. doi: 10.3389/fmicb.2022.771296

Doney, S. C., Ruckelshaus, M., Emmett Duffy, J., Barry, J. P., Chan, F., English, C. A., Galindo, H. M., Grebmeier, J. M., Hollowed, A. B., Knowlton, N., Polovina, J., Rabalais, N. N., Sydeman, W. J., & Talley, L. D. (2012). Climate Change Impacts on Marine Ecosystems. *Annual Review of Marine Science*, *4*(1), 11–37. doi: 10.1146/annurev-marine-041911-111611

Doolittle, W. F. (2012). Population Genomics: How Bacterial Species Form and Why They Don't Exist. *Current Biology*, *22*(11), R451–R453. doi: 10.1016/j.cub.2012.04.034

Drake, L. A., Doblin, M. A., & Dobbs, F. C. (2007). Potential microbial bioinvasions via ships' ballast water, sediment, and biofilm. *Marine Pollution Bulletin*, *55*(7–9), 333–341. doi: 10.1016/j.marpolbul.2006.11.007

Eckert, E. M., Galafassi, S., Bastidas Navarro, M., Di Cesare, A., & Corno, G. (2023). Increased similarity of aquatic bacterial communities of different origin after antibiotic disturbance. *Environmental Pollution*, *316*, 120568. doi: 10.1016/j.envpol.2022.120568

Ferrenberg, S., O'Neill, S. P., Knelman, J. E., Todd, B., Duggan, S., Bradley, D., Robinson, T., Schmidt, S. K., Townsend, A. R., Williams, M. W., Cleveland, C. C., Melbourne, B. A., Jiang, L., & Nemergut, D. R. (2013). Changes in assembly processes in soil bacterial communities following a wildfire disturbance. *The ISME Journal*, *7*(6), 1102–1111. doi: 10.1038/ismej.2013.11

Fillinger, L., Zhou, Y., Kellermann, C., & Griebler, C. (2019). Non-random processes determine the colonization of groundwater sediments by microbial communities in a pristine porous aquifer. *Environmental Microbiology*, *21*(1), 327–342. doi: 10.1111/1462-2920.14463

Fitzgerald, C., & Nachamkin, I. (2015). Campylobacter and Arcobacter. In *Manual of Clinical Microbiology* (pp. 998–1012). ASM Press. doi: 10.1128/9781555817381.ch56

Fjellheim, A. J., Playfoot, K. J., Skjermo, J., & Vadstein, O. (2012). Inter-individual variation in the dominant intestinal microbiota of reared Atlantic cod (Gadus morhua L.) larvae. *Aquaculture Research*, *43*(10), 1499–1508. doi: 10.1111/j.1365-2109.2011.02952.x

Fox, J. W. (2013a). The intermediate disturbance hypothesis is broadly defined, substantive issues are key: a reply to Sheil and Burslem. *Trends in Ecology & Evolution*, *28*(10), 572–573. doi: 10.1016/j.tree.2013.07.008

Fox, J. W. (2013b). The intermediate disturbance hypothesis should be abandoned. *Trends in Ecology & Evolution*, *28*(2), 86–92. doi: 10.1016/j.tree.2012.08.014

Frenk, S., Hadar, Y., & Minz, D. (2018). Quality of Irrigation Water Affects Soil Functionality and Bacterial Community Stability in Response to Heat Disturbance. *Applied and Environmental Microbiology*, *84*(4), 2087–2104. doi: 10.1128/AEM.02087-17

Fuhrman, J. A., & Hagstrm, ke. (2008). Bacterial and Archaeal Community Structure and its Patterns. In *Microbial Ecology of the Oceans* (pp. 45–90). John Wiley & Sons, Inc. doi: 10.1002/9780470281840.ch3

Fukami, T. (2004). Assembly history interacts with ecosystem size to influence species diversity. *Ecology*, *85*(12), 3234–3242. doi: 10.1890/04-0340

Ganeshan, & Hosseinidoust. (2019). Phage Therapy with a focus on the Human Microbiota. *Antibiotics*, *8*(3), 131. doi: 10.3390/antibiotics8030131

Garnier, A., Pennekamp, F., Lemoine, M., & Petchey, O. L. (2017). Temporal scale dependent interactions between multiple environmental disturbances in microcosm ecosystems. *Global Change Biology*, *23*(12), 5237–5248. doi: 10.1111/gcb.13786

Gibbons, S. M., Scholz, M., Hutchison, A. L., Dinner, A. R., Gilbert, J. A., & Coleman, M. L. (2016). Disturbance Regimes Predictably Alter Diversity in an Ecologically Complex Bacterial System. *MBio*, *7*(6). doi: 10.1128/mBio.01372-16

Gill, J., & Hyman, P. (2010). Phage Choice, Isolation, and Preparation for Phage Therapy. *Current Pharmaceutical Biotechnology*, *11*(1), 2–14. doi: 10.2174/138920110790725311

Gotelli, N. J., & Ulrich, W. (2012). Statistical challenges in null model analysis. *Oikos*, *121*(2), 171–180. doi: 10.1111/j.1600-0706.2011.20301.x

Graham, D. W., & Smith, V. H. (2004). Designed Ecosystem Services: Application of Ecological Principles in Wastewater Treatment Engineering. *Frontiers in Ecology and the Environment*, *2*(4), 199. doi: 10.2307/3868315

Grenni, P., Ancona, V., & Barra Caracciolo, A. (2018). Ecological effects of antibiotics on natural ecosystems: A review. *Microchemical Journal*, *136*, 25–39. doi: 10.1016/j.microc.2017.02.006

Griffiths, B. S., & Philippot, L. (2013). Insights into the resistance and resilience of the soil microbial community. *FEMS Microbiology Reviews*, *37*(2), 112–129. doi: 10.1111/j.1574-6976.2012.00343.x

Grubb, D. S., Wrigley, S. D., Freedman, K. E., Wei, Y., Vazquez, A. R., Trotter, R. E., Wallace, T. C., Johnson, S. A., & Weir, T. L. (2020). PHAGE-2 Study: Supplemental Bacteriophages Extend Bifidobacterium animalis subsp. lactis BL04 Benefits on Gut Health and Microbiota in Healthy Adults. *Nutrients*, *12*(8), 2474. doi: 10.3390/nu12082474

Hahn, M. W., & Höfle, M. G. (1999). Flagellate Predation on a Bacterial Model Community: Interplay of Size-Selective Grazing, Specific Bacterial Cell Size, and Bacterial Community Composition. *Applied and Environmental Microbiology*, *65*(11), 4863–4872. doi: 10.1128/AEM.65.11.4863-4872.1999

Hanson, C. A., Fuhrman, J. A., Horner-Devine, M. C., & Martiny, J. B. H. (2012). Beyond biogeographic patterns: processes shaping the microbial landscape. *Nature Reviews Microbiology*, *10*(7), 497–506. doi: 10.1038/nrmicro2795

A Hanson, C., Horner-Devine, M. C., Martiny, J. B. H., & Fuhrman, J. A. (2013). Microbial Biogeography. In *Encyclopedia of Biodiversity* (pp. 271–279). Elsevier. doi: 10.1016/B978-0-12-384719-5.00402-0

Harper, A., Naghibi, M., & Garcha, D. (2018). The Role of Bacteria, Probiotics and Diet in Irritable Bowel Syndrome. *Foods*, *7*(2), 13. doi: 10.3390/foods7020013

Herren, C. M., Webert, K. C., & McMahon, K. D. (2016). Environmental Disturbances Decrease the Variability of Microbial Populations within Periphyton. *MSystems*, *1*(3), 13–16. doi: 10.1128/mSystems.00013-16

Hess-Erga, O.-K., Blomvågnes-Bakke, B., & Vadstein, O. (2010). Recolonization by heterotrophic bacteria after UV irradiation or ozonation of seawater; a simulation of ballast water treatment. *Water Research*, *44*(18), 5439–5449. doi: 10.1016/j.watres.2010.06.059

Heys, C., Cheaib, B., Busetti, A., Kazlauskaite, R., Maier, L., Sloan, W. T., Ijaz, U. Z., Kaufmann, J., McGinnity, P., & Llewellyn, M. S. (2020). Neutral Processes Dominate Microbial Community Assembly in Atlantic Salmon, *Salmo salar*. *Applied and Environmental Microbiology*, *86*(8). doi: 10.1128/AEM.02283-19

Higuera, G., Bastías, R., Tsertsvadze, G., Romero, J., & Espejo, R. T. (2013). Recently discovered Vibrio anguillarum phages can protect against experimentally induced vibriosis in Atlantic salmon, Salmo salar. *Aquaculture*, *392–395*, 128–133. doi: 10.1016/j.aquaculture.2013.02.013

Himeoka, Y., & Mitarai, N. (2020). Dynamics of bacterial populations under the feast-famine cycles. *Physical Review Research*, *2*(1), 013372. doi: 10.1103/PhysRevResearch.2.013372

Hong, Y., Thimmapuram, J., Zhang, J., Collings, C. K., Bhide, K., Schmidt, K., & Ebner, P. D. (2016). The impact of orally administered phages on host immune response and surrounding microbial communities. *Bacteriophage*, *6*(3), e1211066. doi: 10.1080/21597081.2016.1211066

Hsu, B. B., Gibson, T. E., Yeliseyev, V., Liu, Q., Lyon, L., Bry, L., Silver, P. A., & Gerber, G. K. (2019). Dynamic Modulation of the Gut Microbiota and Metabolome by Bacteriophages in a Mouse Model. *Cell Host & Microbe*, *25*(6), 803-814.e5. doi: 10.1016/j.chom.2019.05.001

Hu, J., Amor, D. R., Barbier, M., Bunin, G., & Gore, J. (2022). Emergent phases of ecological diversity and dynamics mapped in microcosms. *Science*, *378*(6615), 85–89. doi: 10.1126/science.abm7841

Hu, Y. O. O., Hugerth, L. W., Bengtsson, C., Alisjahbana, A., Seifert, M., Kamal, A., Sjöling, Å., Midtvedt, T., Norin, E., Du, J., & Engstrand, L. (2018). Bacteriophages Synergize with the Gut Microbial Community To Combat *Salmonella*. *MSystems*, *3*(5). doi: 10.1128/mSystems.00119-18

Hubbell, S. P. (2001). *The unified neutral theory of biodiversity and biogeography*. Princeton University Press.

Hubbell, S. (2005). Neutral theory in community ecology and the hypothesis of functional equivalence. *Functional Ecology*, *19*(1), 166–172. doi: 10.1111/j.0269-8463.2005.00965.x

Hutchins, D. A., & Fu, F. (2017). Microorganisms and ocean global change. *Nature Microbiology*, *2*(6), 17058. doi: 10.1038/nmicrobiol.2017.58

Hutchinson, G. E. (1961). The Paradox of the Plankton. *The American Naturalist*, *95*(882), 137–145. http://www.jstor.org/stable/2458386

Hutchinson, E. (1965). The Ecological Theatre and the Evolutionary Play. In *BioScience* (Issue 8). Yale University Press. doi: 10.2307/1293709

Jiao, S., Yang, Y., Xu, Y., Zhang, J., & Lu, Y. (2020). Balance between community assembly processes mediates species coexistence in agricultural soil microbiomes across eastern China. *The ISME Journal*, *14*(1), 202–216. doi: 10.1038/s41396-019-0522-9

Jo, J.-H., Kennedy, E. A., & Kong, H. H. (2016). Research Techniques Made Simple: Bacterial 16S Ribosomal RNA Gene Sequencing in Cutaneous Research. *Journal of Investigative Dermatology*, *136*(3), e23–e27. doi: 10.1016/j.jid.2016.01.005

Johnke, J., Cohen, Y., de Leeuw, M., Kushmaro, A., Jurkevitch, E., & Chatzinotas, A. (2014). Multiple micro-predators controlling bacterial communities in the environment. *Current Opinion in Biotechnology*, *27*, 185–190. doi: 10.1016/j.copbio.2014.02.003

Pantoja-Don Juan, C. A., Gómez-Verduzco, G., Márquez-Mota, C. C., Téllez-Isaías, G., Kwon, Y. M., Cortés-Cuevas, A., Arce-Menocal, J., Martínez-Gómez, D., & Ávila-González, E. (2022). Productive Performance and Cecum Microbiota Analysis of Broiler Chickens Supplemented with β-Mannanases and Bacteriophages—A Pilot Study. *Animals*, *12*(2), 169. doi: 10.3390/ani12020169

Kalatzis, P. G., Bastías, R., Kokkari, C., & Katharios, P. (2016). Isolation and Characterization of Two Lytic Bacteriophages, ϕSt2 and ϕGrn1; Phage Therapy Application for Biological Control of Vibrio alginolyticus in Aquaculture Live Feeds. *PLOS ONE*, *11*(3), e0151101. doi: 10.1371/journal.pone.0151101

Karunasagar, I., Shivu, M. M., Girisha, S. K., Krohne, G., & Karunasagar, I. (2007). Biocontrol of pathogens in shrimp hatcheries using bacteriophages. *Aquaculture*, *268*(1–4), 288–292. doi: 10.1016/j.aquaculture.2007.04.049

Kearns, P. J., & Shade, A. (2018). Trait-based patterns of microbial dynamics in dormancy potential and heterotrophic strategy: case studies of resource-based and post-press succession. *The ISME Journal*, *12*(11), 2575–2581. doi: 10.1038/s41396-018-0194-x

Keating, C., Bolton-Warberg, M., Hinchcliffe, J., Davies, R., Whelan, S., Wan, A. H. L., Fitzgerald, R. D., Davies, S. J., Smith, C. J., & Ijaz, U. Z. (2022). Drivers of ecological assembly in the hindgut of Atlantic Cod fed a macroalgal supplemented diet. *Npj Biofilms and Microbiomes*, *8*(1), 36. doi: 10.1038/s41522-022-00296-x

Keesing, F., Belden, L. K., Daszak, P., Dobson, A., Harvell, C. D., Holt, R. D., Hudson, P., Jolles, A., Jones, K. E., Mitchell, C. E., Myers, S. S., Bogich, T., & Ostfeld, R. S. (2010). Impacts of biodiversity on the emergence and transmission of infectious diseases. *Nature*, *468*(7324), 647–652. doi: 10.1038/nature09575

Kinnunen, M., Dechesne, A., Albrechtsen, H.-J., & Smets, B. F. (2018). Stochastic processes govern invasion success in microbial communities when the invader is phylogenetically close to resident bacteria. *The ISME Journal*, *12*(11), 2748–2756. doi: 10.1038/s41396-018-0202-1

Kinnunen, M., Dechesne, A., Proctor, C., Hammes, F., Johnson, D., Quintela-Baluja, M., Graham, D., Daffonchio, D., Fodelianakis, S., Hahn, N., Boon, N., & Smets, B. F. (2016). A conceptual framework for invasion in microbial communities. *The ISME Journal*, *10*(12), 2773–2779. doi: 10.1038/ismej.2016.75

Kirchman, D. L. (2012). *Processes in microbial ecology*. Oxford University Press.

Klappenbach, J. A., Dunbar, J. M., & Schmidt, T. M. (2000). rRNA Operon Copy Number Reflects Ecological Strategies of Bacteria. *Applied and Environmental Microbiology*, *66*(4), 1328–1333. doi: 10.1128/AEM.66.4.1328-1333.2000

Konopka, A. (2009). What is microbial community ecology? *The ISME Journal*, *3*(11), 1223–1230. doi: 10.1038/ismej.2009.88

Lear, L., Padfield, D., Inamine, H., Shea, K., & Buckling, A. (2022). Disturbance-mediated invasions are dependent on community resource abundance. *Ecology*, *103*(8), e3728. doi: 10.1002/ecy.3728

Leeuwenhoek, A. (1677). Observations, communicated to the publisher by Mr. Antony van Leewenhoeck, in a dutch letter of the 9th Octob. 1676. here English'd: concerning little animals by him observed in rain-well-sea- and snow water; as also in water wherein pepper had lain infused. *Philosophical Transactions of the Royal Society of London*, *12*(133), 821–831. doi: 10.1098/rstl.1677.0003

Legendre, P., Borcard, D., & Peres-Neto, P. R. (2005). Analyzing beta diversity: partitioning the spatial variation of community composition data. *Ecological Monographs*, *75*(4), 435–450. doi: 10.1890/05-0549

Leibold, M. A., Holyoak, M., Mouquet, N., Amarasekare, P., Chase, J. M., Hoopes, M. F., Holt, R. D., Shurin, J. B., Law, R., Tilman, D., Loreau, M., & Gonzalez, A. (2004). The metacommunity concept: a framework for multi-scale community ecology. *Ecology Letters*, *7*(7), 601–613. doi: 10.1111/j.1461-0248.2004.00608.x

Lewis, S. L., & Maslin, M. A. (2015). Defining the Anthropocene. *Nature*, *519*(7542), 171–180. doi: 10.1038/nature14258

Li, X., Zhou, L., Yu, Y., Ni, J., Xu, W., & Yan, Q. (2017). Composition of Gut Microbiota in the Gibel Carp (Carassius auratus gibelio) Varies with Host Development. *Microbial Ecology*, *74*(1), 239–249. doi: 10.1007/s00248-016-0924-4

Litchman, E. (2010). Invisible invaders: non-pathogenic invasive microbes in aquatic and terrestrial ecosystems. *Ecology Letters*, *13*(12), 1560–1572. doi: 10.1111/j.1461-0248.2010.01544.x

Liu, M., Bjørnlund, L., Rønn, R., Christensen, S., & Ekelund, F. (2012). Disturbance Promotes Non-Indigenous Bacterial Invasion in Soil Microcosms: Analysis of the Roles of Resource Availability and Community Structure. *PLoS ONE*, *7*(10), e45306. doi: 10.1371/journal.pone.0045306

Liu, S., Wawrik, B., & Liu, Z. (2017). Different Bacterial Communities Involved in Peptide Decomposition between Normoxic and Hypoxic Coastal Waters. *Frontiers in Microbiology*, *8*(MAR), 353. doi: 10.3389/fmicb.2017.00353

Liu, W., Graham, E. B., Dong, Y., Zhong, L., Zhang, J., Qiu, C., Chen, R., Lin, X., & Feng, Y. (2021). Balanced stochastic versus deterministic assembly processes benefit diverse yet uneven ecosystem functions in representative agroecosystems. *Environmental Microbiology*, *23*(1), 391–404. doi: 10.1111/1462-2920.15326

Locey, K. J., & Lennon, J. T. (2019). A Residence Time Theory for Biodiversity. *The American Naturalist*, *194*(1), 59–72. doi: 10.1086/703456

Low-Décarie, E., Kolber, M., Homme, P., Lofano, A., Dumbrell, A., Gonzalez, A., & Bell, G. (2015). Community rescue in experimental metacommunities. *PNAS*, *112*(46), 14307–14312. doi: 10.1073/pnas.1513125112

Lucas, R., Groeneveld, J., Harms, H., Johst, K., Frank, K., & Kleinsteuber, S. (2017). A critical evaluation of ecological indices for the comparative analysis of microbial communities based on molecular datasets. *FEMS Microbiology Ecology*, *93*(1), fiw209. doi: 10.1093/femsec/fiw209

Lulijwa, R., Rupia, E. J., & Alfaro, A. C. (2020). Antibiotic use in aquaculture, policies and regulation, health and environmental risks: a review of the top 15 major producers. *Reviews in Aquaculture*, *12*(2), 640–663. doi: 10.1111/raq.12344

MacArthur, R. H., & Wilson, O. E. (1967). *The theory of island biogeography* (Vol. 1). Princeton University Press.

Mallon, C. A., Elsas, J. D. van, & Salles, J. F. (2015). Microbial Invasions: The Process, Patterns, and Mechanisms. *Trends in Microbiology*, *23*(11), 719–729. doi: 10.1016/j.tim.2015.07.013

Mallon, C. A., Le Roux, X., van Doorn, G. S., Dini-Andreote, F., Poly, F., & Salles, J. F. (2018). The impact of failure: unsuccessful bacterial invasions steer the soil microbial community away from the invader's niche. *The ISME Journal*, *12*(3), 728–741. doi: 10.1038/s41396-017-0003-y

Mannina, G., Gulhan, H., & Ni, B.-J. (2022). Water reuse from wastewater treatment: The transition towards circular economy in the water sector. *Bioresource Technology*, *363*, 127951. doi: 10.1016/j.biortech.2022.127951

Mao, Z., Zhao, Z., Da, J., Tao, Y., Li, H., Zhao, B., Xing, P., & Wu, Q. (2023). The selection of copiotrophs may complicate biodiversity-ecosystem functioning relationships in microbial dilution-to-extinction experiments. *Environmental Microbiome*, *18*(1), 19. doi: 10.1186/s40793-023-00478-w

May, R. M. (1974). *Stability and Complexity in Model Ecosystems* (Vol. 1). Princeton University Press. http://www.jstor.org/stable/10.2307/j.ctvs32rq4

McCallin, S., Alam Sarker, S., Barretto, C., Sultana, S., Berger, B., Huq, S., Krause, L., Bibiloni, R., Schmitt, B., Reuteler, G., & Brüssow, H. (2013). Safety analysis of a Russian phage cocktail: From MetaGenomic analysis to oral application in healthy human subjects. *Virology*, *443*(2), 187–196. doi: 10.1016/j.virol.2013.05.022

Münkemüller, T., Lavergne, S., Bzeznik, B., Dray, S., Jombart, T., Schiffers, K., & Thuiller, W. (2012). How to measure and test phylogenetic signal. *Methods in Ecology and Evolution*, *3*(4), 743–756. doi: 10.1111/j.2041-210X.2012.00196.x

Navarrete, P., Mardones, P., Opazo, R., Espejo, R., & Romero, J. (2008). Oxytetracycline Treatment Reduces Bacterial Diversity of Intestinal Microbiota of Atlantic Salmon. *Journal of Aquatic Animal Health*, *20*(3), 177–183. doi: 10.1577/H07-043.1

Nemergut, D. R., Knelman, J. E., Ferrenberg, S., Bilinski, T., Melbourne, B., Jiang, L., Violle, C., Darcy, J. L., Prest, T., Schmidt, S. K., & Townsend, A. R. (2016). Decreases in average bacterial community rRNA operon copy number during succession. *The ISME Journal*, *10*(5), 1147–1156. doi: 10.1038/ismej.2015.191

Nemergut, D. R., Schmidt, S. K., Fukami, T., O'Neill, S. P., Bilinski, T. M., Stanish, L. F., Knelman, J. E., Darcy, J. L., Lynch, R. C., Wickey, P., & Ferrenberg, S. (2013). Patterns and Processes of Microbial Community Assembly. *Microbiology and Molecular Biology Reviews*, *77*(3), 342–356. doi: 10.1128/MMBR.00051-12

Newton, D. P., Ho, P.-Y., & Huang, K. C. (2023). Modulation of antibiotic effects on microbial communities by resource competition. *Nature Communications*, *14*(1), 2398. doi: 10.1038/s41467-023-37895-x

Ning, D., Deng, Y., Tiedje, J. M., & Zhou, J. (2019). A general framework for quantitatively assessing ecological stochasticity. *PNAS*, *116*(34), 16892–16898. doi: 10.1073/pnas.1904623116

Ning, D., Yuan, M., Wu, L., Zhang, Y., Guo, X., Zhou, X., Yang, Y., Arkin, A. P., Firestone, M. K., & Zhou, J. (2020). A quantitative framework reveals ecological drivers of grassland microbial community assembly in response to warming. *Nature Communications*, *11*(1), 4717. doi: 10.1038/s41467-020-18560-z

Noble, R., & Fuhrman, J. (1999). Breakdown and microbial uptake of marine viruses and other lysis products. *Aquatic Microbial Ecology*, *20*(1), 1–11. doi: 10.3354/ame020001

Noble, R., Middelboe, M., & Fuhrman, J. (1999). Effects of viral enrichment on the mortality and growth of heterotrophic bacterioplankton. *Aquatic Microbial Ecology*, *18*(1), 1–13. doi: 10.3354/ame018001

Odum, E. P. (1971). *Fundamentals of ecology* (3rd ed.). W. B. Saunders.

Oliveira, J., Castilho, F., Cunha, A., & Pereira, M. J. (2012). Bacteriophage therapy as a bacterial control strategy in aquaculture. *Aquaculture International*, *20*(5), 879–910. doi: 10.1007/s10499-012-9515-7

Osman, R. W. (2015). The Intermediate Disturbance Hypothesis. In *Encyclopedia of Ecology* (pp. 441–450). Elsevier. doi: 10.1016/B978-0-12-409548-9.09480-X

Ovaskainen, O., Tikhonov, G., Dunson, D., Grøtan, V., Engen, S., Sæther, B.-E., & Abrego, N. (2017). How are species interactions structured in species-rich communities? A new method for analysing time-series data. *Proceedings of the Royal Society B: Biological Sciences*, *284*(1855), 20170768. doi: 10.1098/rspb.2017.0768

Paine, R. T. (1969). A Note on Trophic Complexity and Community Stability. *The American Naturalist*, *103*(929), 91–93. http://www.jstor.org/stable/2459472

Pallecchi, L., Bartoloni, A., Paradisi, F., & Rossolini, G. M. (2008). Antibiotic resistance in the absence of antimicrobial use: mechanisms and implications. *Expert Review of Anti-Infective Therapy*, *6*(5), 725–732. doi: 10.1586/14787210.6.5.725

Patangia, D. V., Anthony Ryan, C., Dempsey, E., Paul Ross, R., & Stanton, C. (2022). Impact of antibiotics on the human microbiome and consequences for host health. *MicrobiologyOpen*, *11*(1), e1260. doi: 10.1002/mbo3.1260

Peix, A., Ramírez-Bahena, M.-H., & Velázquez, E. (2009). Historical evolution and current status of the taxonomy of genus Pseudomonas. *Infection, Genetics and Evolution*, *9*(6), 1132–1147. doi: 10.1016/j.meegid.2009.08.001

Pereira, C., Silva, Y. J., Santos, A. L., Cunha, Â., Gomes, N. C. M., & Almeida, A. (2011). Bacteriophages with Potential for Inactivation of Fish Pathogenic Bacteria: Survival, Host Specificity and Effect on Bacterial Community Structure. *Marine Drugs*, *9*(11), 2236–2255. doi: 10.3390/md9112236

Perry, W. B., Lindsay, E., Payne, C. J., Brodie, C., & Kazlauskaite, R. (2020). The role of the gut microbiome in sustainable teleost aquaculture. *Proceedings of the Royal Society B: Biological Sciences*, *287*(1926). doi: 10.1098/rspb.2020.0184

Philippot, L., Griffiths, B. S., & Langenheder, S. (2021). Microbial Community Resilience across Ecosystems and Multiple Disturbances. *Microbiology and Molecular Biology Reviews*, *85*(2). doi: 10.1128/MMBR.00026-20

Pörtner, H.-O., Roberts, D. C., Tignor, M., Poloczanska, E. S., Mintenbeck, K., Alegría, A., Craig, M., Langsdorf, S., Löschke, S., Möller, V., Okem, A., & Rama, B. (2022). *IPCC Climate Change 2022: Impacts, Adaptation, and Vulnerability*.

Pramanik, S. K., Suja, F. B., Zain, S. M., & Pramanik, B. K. (2019). The anaerobic digestion process of biogas production from food waste: Prospects and constraints. *Bioresource Technology Reports*, *8*, 100310. doi: 10.1016/j.biteb.2019.100310

Proctor, L. M., & Fuhrman, J. A. (1990). Viral mortality of marine bacteria and cyanobacteria. *Nature*, *343*(6253), 60–62. doi: 10.1038/343060a0

Pullerits, K., Ahlinder, J., Holmer, L., Salomonsson, E., Öhrman, C., Jacobsson, K., Dryselius, R., Forsman, M., Paul, C. J., & Rådström, P. (2020). Impact of UV irradiation at full scale on bacterial communities in drinking water. *Npj Clean Water*, *3*(1), 11. doi: 10.1038/s41545-020-0057-7

Rachid, R., & Chatila, T. A. (2016). The role of the gut microbiota in food allergy. *Current Opinion in Pediatrics*, *28*(6), 748–753. doi: 10.1097/MOP.0000000000000427

Ramakrishna, W., Yadav, R., & Li, K. (2019). Plant growth promoting bacteria in agriculture: Two sides of a coin. *Applied Soil Ecology*, *138*, 10–18. doi: 10.1016/j.apsoil.2019.02.019

Reichenbach, H. (2006). The Order Cytophagales. In *The Prokaryotes* (pp. 549–590). Springer New York. doi: 10.1007/0-387-30747-8_20

Richards, P. J., Connerton, P. L., & Connerton, I. F. (2019). Phage Biocontrol of Campylobacter jejuni in Chickens Does Not Produce Collateral Effects on the Gut Microbiota. *Frontiers in Microbiology*, *10*(MAR), 476. doi: 10.3389/fmicb.2019.00476

Rivett, D. W., Jones, M. L., Ramoneda, J., Mombrikotb, S. B., Ransome, E., & Bell, T. (2018). Elevated success of multispecies bacterial invasions impacts community composition during ecological succession. *Ecology Letters*, *21*(4), 516–524. doi: 10.1111/ele.12916

Roller, B. R. K., Stoddard, S. F., & Schmidt, T. M. (2016). Exploiting rRNA operon copy number to investigate bacterial reproductive strategies. *Nature Microbiology*, *1*(11), 16160. doi: 10.1038/nmicrobiol.2016.160

Rosas-Plaza, S., Hernández-Terán, A., Navarro-Díaz, M., Escalante, A. E., Morales-Espinosa, R., & Cerritos, R. (2022). Human Gut Microbiome Across Different Lifestyles: From Hunter-Gatherers to Urban Populations. *Frontiers in Microbiology*, *13*, 868. doi: 10.3389/fmicb.2022.843170

Rosindell, J., Hubbell, S. P., & Etienne, R. S. (2011). The Unified Neutral Theory of Biodiversity and Biogeography at Age Ten. *Trends in Ecology & Evolution*, *26*(7), 340–348. doi: 10.1016/j.tree.2011.03.024

Salmond, G. P. C., & Fineran, P. C. (2015). A century of the phage: past, present and future. *Nature Reviews Microbiology*, *13*(12), 777–786. doi: 10.1038/nrmicro3564

Salvesen, I., Skjermo, J., & Vadstein, O. (1999). Growth of turbot (Scophthalmus maximus L.) during first feeding in relation to the proportion of r/K-strategists in the bacterial community of the rearing water. *Aquaculture*, *175*(3–4), 337–350. doi: 10.1016/S0044-8486(99)00110-6

Santillan, E., Constancias, F., & Wuertz, S. (2020). Press Disturbance Alters Community Structure and Assembly Mechanisms of Bacterial Taxa and Functional Genes in Mesocosm-Scale Bioreactors. *MSystems*, *5*(4). doi: 10.1128/mSystems.00471-20

Santillan, E., Seshan, H., Constancias, F., Drautz-Moses, D. I., & Wuertz, S. (2019). Frequency of disturbance alters diversity, function, and underlying assembly mechanisms of complex bacterial communities. *Npj Biofilms and Microbiomes*, *5*(1), 8. doi: 10.1038/s41522-019-0079-4

Santillan, E., & Wuertz, S. (2022). Microbiome assembly predictably shapes diversity across a range of disturbance frequencies in experimental microcosms. *Npj Biofilms and Microbiomes*, *8*(1), 41. doi: 10.1038/s41522-022-00301-3

Sarker, S. A., Berger, B., Deng, Y., Kieser, S., Foata, F., Moine, D., Descombes, P., Sultana, S., Huq, S., Bardhan, P. K., Vuillet, V., Praplan, F., & Brüssow, H. (2017). Oral application of Escherichia coli bacteriophage: safety tests in healthy and diarrheal children from Bangladesh. *Environmental Microbiology*, *19*(1), 237–250. doi: 10.1111/1462-2920.13574

Scheffer, M., Carpenter, S., Foley, J. A., Folke, C., & Walker, B. (2001). Catastrophic shifts in ecosystems. *Nature*, *413*(6856), 591–596. doi: 10.1038/35098000

Shade, A. (2017). Diversity is the question, not the answer. *The ISME Journal*, *11*(1), 1–6. doi: 10.1038/ismej.2016.118

Shade, A. (2023). Microbiome rescue: directing resilience of environmental microbial communities. *Current Opinion in Microbiology*, *72*, 102263. doi: 10.1016/j.mib.2022.102263

Shade, A., Gregory Caporaso, J., Handelsman, J., Knight, R., & Fierer, N. (2013). A meta-analysis of changes in bacterial and archaeal communities with time. *The ISME Journal*, *7*(8), 1493–1506. doi: 10.1038/ismej.2013.54

Shade, A., Peter, H., Allison, S. D., Baho, D. L., Berga, M., Bürgmann, H., Huber, D. H., Langenheder, S., Lennon, J. T., Martiny, J. B. H., Matulich, K. L., Schmidt, T. M., & Handelsman, J. (2012). Fundamentals of Microbial Community Resistance and Resilience. *Frontiers in Microbiology*, *3*. doi: 10.3389/fmicb.2012.00417

Sheil, D., & Burslem, D. F. R. P. (2013). Defining and defending Connell's intermediate disturbance hypothesis: a response to Fox. *Trends in Ecology & Evolution*, *28*(10), 571–572. doi: 10.1016/j.tree.2013.07.006

Shen, Y., Ryser, E. T., Li, H., & Zhang, W. (2021). Bacterial community assembly and antibiotic resistance genes in the lettuce-soil system upon antibiotic exposure. *Science of The Total Environment*, *778*, 146255. doi: 10.1016/j.scitotenv.2021.146255

Shi, Y., Li, Y., Xiang, X., Sun, R., Yang, T., He, D., Zhang, K., Ni, Y., Zhu, Y.-G., Adams, J. M., & Chu, H. (2018). Spatial scale affects the relative role of stochasticity versus determinism in soil bacterial communities in wheat fields across the North China Plain. *Microbiome*, *6*(1), 27. doi: 10.1186/s40168-018-0409-4

Skjermo, J., Salvesen, I., Øie, G., Olsen, Y., & Vadstein, O. (1997). Microbially matured water: a technique for selection of a non-opportunistic bacterial flora in water that may improve performance of marine larvae. *Aquaculture International*, *5*(1), 13–28. doi: 10.1007/BF02764784

Sloan, W. T., Lunn, M., Woodcock, S., Head, I. M., Nee, S., & Curtis, T. P. (2006). Quantifying the roles of immigration and chance in shaping prokaryote community structure. *Environmental Microbiology*, *8*(4), 732–740. doi: 10.1111/j.1462-2920.2005.00956.x

Sorensen, J. W., & Shade, A. (2020). Dormancy dynamics and dispersal contribute to soil microbiome resilience. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *375*(1798), 20190255. doi: 10.1098/rstb.2019.0255

Stegen, J. C., Lin, X., Fredrickson, J. K., Chen, X., Kennedy, D. W., Murray, C. J., Rockhold, M. L., & Konopka, A. (2013). Quantifying community assembly processes and identifying features that impose them. *The ISME Journal*, *7*(11), 2069–2079. doi: 10.1038/ismej.2013.93

Stegen, J. C., Lin, X., Konopka, A. E., & Fredrickson, J. K. (2012). Stochastic and deterministic assembly processes in subsurface microbial communities. *The ISME Journal*, *6*(9), 1653–1664. doi: 10.1038/ismej.2012.22

Sullam, K. E., & Musa, T. (2021). Ecological Dynamics and Microbial Treatments against Oomycete Plant Pathogens. *Plants*, *10*(12), 2697. doi: 10.3390/plants10122697

Theron, J., & Cloete, T. E. (2000). Molecular Techniques for Determining Microbial Diversity and Community Structure in Natural Environments. *Critical Reviews in Microbiology*, *26*(1), 37–57. doi: 10.1080/10408410091154174

Thingstad, T. F. (2000). Elements of a theory for the mechanisms controlling abundance, diversity, and biogeochemical role of lytic bacterial viruses in aquatic systems. *Limnology and Oceanography*, *45*(6), 1320–1328. doi: 10.4319/lo.2000.45.6.1320

Thiruselvi, D., Kumar, P. S., Kumar, M. A., Lay, C.-H., Aathika, S., Mani, Y., Jagadiswary, D., Dhanasekaran, A., Shanmugam, P., Sivanesan, S., & Show, P.-L. (2021). A critical review on global trends in biogas scenario with its up-gradation techniques for fuel cell and future perspectives. *International Journal of Hydrogen Energy*, *46*(31), 16734–16750. doi: 10.1016/j.ijhydene.2020.10.023

Tram VO, P., Ngo, H. H., Guo, W., Zhou, J. L., Nguyen, P. D., Listowski, A., & Wang, X. C. (2014). A mini-review on the impacts of climate change on wastewater reclamation and reuse. *Science of The Total Environment*, *494–495*, 9–17. doi: 10.1016/j.scitotenv.2014.06.090

Tuomisto, H. (2010). A diversity of beta diversities: straightening up a concept gone awry. Part 1. Defining beta diversity as a function of alpha and gamma diversity. *Ecography*, *33*(1), 2–22. doi: 10.1111/j.1600-0587.2009.05880.x

Tuomisto, H. (2011). Commentary: do we have a consistent terminology for species diversity? Yes, if we choose to use it. *Oecologia*, *167*(4), 903–911. doi: 10.1007/s00442-011-2128-4

Unger, I. M., Goyne, K. W., Kennedy, A. C., Kremer, R. J., McLain, J. E. T., & Williams, C. F. (2013). Antibiotic Effects on Microbial Community Characteristics in Soils under Conservation Management Practices. *Soil Science Society of America Journal*, *77*(1), 100–112. doi: 10.2136/sssaj2012.0099

United Nations. (2022). *World Population Prospects 2022 -Summary of Results*.

United Nations. (2023). *The 17 goals - Sustainable Development*. https://sdgs.un.org/goals

Vadstein, O., Attramadal, K. J. K., Bakke, I., Forberg, T., Olsen, Y., Verdegem, M., Giatsis, C., Skjermo, J., Aasen, I. M., Gatesoupe, F.-J., Dierckens, K., Sorgeloos, P., & Bossier, P. (2018). Managing the Microbial Community of Marine Fish Larvae: A Holistic Perspective for Larviculture. *Frontiers in Microbiology*, *9*, 1820. doi: 10.3389/fmicb.2018.01820

Vadstein, O., Attramadal, K. J. K., Bakke, I., & Olsen, Y. (2018). K-Selection as Microbial Community Management Strategy: A Method for Improved Viability of Larvae in Aquaculture. *Frontiers in Microbiology*, *9*, 2730. doi: 10.3389/fmicb.2018.02730

van Nes, E. H., & Scheffer, M. (2007). Slow Recovery from Perturbations as a Generic Indicator of a Nearby Catastrophic Shift. *The American Naturalist*, *169*(6), 738–747. doi: 10.1086/516845

Vellend, M. (2010). Conceptual Synthesis in Community Ecology. *The Quarterly Review of Biology*, *85*(2), 183–206. doi: 10.1086/652373

Vellend, M. (2016). *The Theory of Ecological Communities* (Vol. 57). Princeton University Press. https://press.princeton.edu/books/hardcover/9780691164847/the-theory-of-ecological-communities-mpb-57

Vellend, M., Srivastava, D. S., Anderson, K. M., Brown, C. D., Jankowski, J. E., Kleynhans, E. J., Kraft, N. J. B., Letaw, A. D., Macdonald, A. A. M., Maclean, J. E., Myers-Smith, I. H., Norris, A. R., & Xue, X. (2014). Assessing the relative importance of neutral stochasticity in ecological communities. *Oikos*, *123*(12), 1420–1430. doi: 10.1111/oik.01493

Verner-Jeffreys, D. W., Shields, R. J., Bricknell, I. R., & Birkbeck, T. H. (2004). Effects of different water treatment methods and antibiotic addition on larval survival and gut microflora development in Atlantic halibut (Hippoglossus hippoglossus L.) yolk-sac larvae. *Aquaculture*, *232*(1–4), 129–143. doi: 10.1016/S0044-8486(03)00525-8

Vila, J. C. C., Jones, M. L., Patel, M., Bell, T., & Rosindell, J. (2019). Uncovering the rules of microbial community invasions. *Nature Ecology & Evolution*, *3*(8), 1162–1171. doi: 10.1038/s41559-019-0952-9

Vuono, D. C., Benecke, J., Henkel, J., Navidi, W. C., Cath, T. Y., Munakata-Marr, J., Spear, J. R., & Drewes, J. E. (2015). Disturbance and temporal partitioning of the activated sludge metacommunity. *The ISME Journal*, *9*(2), 425–435. doi: 10.1038/ismej.2014.139

Wang, K., Yan, H., Peng, X., Hu, H., Zhang, H., Hou, D., Chen, W., Qian, P., Liu, J., Cai, J., Chai, X., & Zhang, D. (2020). Community assembly of bacteria and archaea in coastal waters governed by contrasting mechanisms: A seasonal perspective. *Molecular Ecology*, *29*(19), 3762–3776. doi: 10.1111/mec.15600

Wennekes, P. L., Rosindell, J., & Etienne, R. S. (2012). The Neutral—Niche Debate: A Philosophical Perspective. *Acta Biotheoretica*, *60*(3), 257–271. doi: 10.1007/s10441-012-9144-6

WHO. (2022). *Global Antimicrobial Resistance and Use Surveillance System (GLASS) Report 2022*.

de Wit, R., & Bouvier, T. (2006). 'Everything is everywhere, but, the environment selects'; what did Baas Becking and Beijerinck really say? *Environmental Microbiology*, *8*(4), 755–758. doi: 10.1111/j.1462-2920.2006.01017.x

Woese, C. R. (1987). Bacterial evolution. *Microbiological Reviews*, *51*(2), 221–271. doi: 10.1128/mr.51.2.221-271.1987

Wu, F., Ha, Y., Weiss, A., Wang, M., Letourneau, J., Wang, S., Luo, N., Huang, S., Lee, C. T., David, L. A., & You, L. (2022). Modulation of microbial community dynamics by spatial partitioning. *Nature Chemical Biology*, *18*(4), 394–402. doi: 10.1038/s41589-021-00961-w

Xiao, F., Zhu, W., Yu, Y., He, Z., Wu, B., Wang, C., Shu, L., Li, X., Yin, H., Wang, J., Juneau, P., Zheng, X., Wu, Y., Li, J., Chen, X., Hou, D., Huang, Z., He, J., Xu, G., … Yan, Q. (2021). Host development overwhelms environmental dispersal in governing the ecological succession of zebrafish gut microbiota. *Npj Biofilms and Microbiomes*, *7*(1), 5. doi: 10.1038/s41522-020-00176-2

Yachi, S., & Loreau, M. (1999). Biodiversity and ecosystem productivity in a fluctuating environment: The insurance hypothesis. *PNAS*, *96*(4), 1463–1468. doi: 10.1073/pnas.96.4.1463

Yan, Q., Li, J., Yu, Y., Wang, J., He, Z., Van Nostrand, J. D., Kempher, M. L., Wu, L., Wang, Y., Liao, L., Li, X., Wu, S., Ni, J., Wang, C., & Zhou, J. (2016). Environmental filtering decreases with fish development for the assembly of gut microbiota. *Environmental Microbiology*, *18*(12), 4739–4754. doi: 10.1111/1462-2920.13365

Yang, W., Zheng, C., Zheng, Z., Wei, Y., Lu, K., & Zhu, J. (2018). Nutrient enrichment during shrimp cultivation alters bacterioplankton assemblies and destroys community stability. *Ecotoxicology and Environmental Safety*, *156*, 366–374. doi: 10.1016/j.ecoenv.2018.03.043

Yassour, M., Vatanen, T., Siljander, H., Hämäläinen, A.-M., Härkönen, T., Ryhänen, S. J., Franzosa, E. A., Vlamakis, H., Huttenhower, C., Gevers, D., Lander, E. S., Knip, M., & Xavier, R. J. (2016). Natural history of the infant gut microbiome and impact of antibiotic treatment on bacterial strain diversity and stability. *Science Translational Medicine*, *8*(343). doi: 10.1126/scitranslmed.aad0917

Yin, Q., Sun, Y., Li, B., Feng, Z., & Wu, G. (2022). The r/K selection theory and its application in biological wastewater treatment processes. *Science of The Total Environment*, *824*, 153836. doi: 10.1016/j.scitotenv.2022.153836

Zhang, F.-G., & Zhang, Q.-G. (2015). Patterns in Species Persistence and Biomass Production in Soil Microcosms Recovering from a Disturbance Reject a Neutral Hypothesis for Bacterial Community Assembly. *PLOS ONE*, *10*(5), e0126962. doi: 10.1371/journal.pone.0126962

Zhang, Y.-J., Li, S., Gan, R.-Y., Zhou, T., Xu, D.-P., & Li, H.-B. (2015). Impacts of Gut Bacteria on Human Health and Diseases. *International Journal of Molecular Sciences*, *16*(12), 7493–7519. doi: 10.3390/ijms16047493

Zhao, H., Li, Y., Lv, P., Huang, J., Tai, R., Jin, X., Wang, J., & Wang, X. (2022). Salmonella Phages Affect the Intestinal Barrier in Chicks by Altering the Composition of Early Intestinal Flora: Association With Time of Phage Use. *Frontiers in Microbiology*, *13*, 2489. doi: 10.3389/fmicb.2022.947640

Zhao, J., Liu, Y., Xiao, C., He, S., Yao, H., & Bao, G. (2017). Efficacy of Phage Therapy in Controlling Rabbit Colibacillosis and Changes in Cecal Microbiota. *Frontiers in Microbiology*, *8*(MAY), 957. doi: 10.3389/fmicb.2017.00957

Zhou, J., Deng, Y., Zhang, P., Xue, K., Liang, Y., Van Nostrand, J. D., Yang, Y., He, Z., Wu, L., Stahl, D. A., Hazen, T. C., Tiedje, J. M., & Arkin, A. P. (2014). Stochasticity, succession, and environmental perturbations in a fluidic ecosystem. *PNAS*, *111*(9), E836–E845. doi: 10.1073/pnas.1324044111

Zhou, J., & Ning, D. (2017). Stochastic Community Assembly: Does It Matter in Microbial Ecology? *Microbiology and Molecular Biology Reviews*, *81*(4), e00002-17. doi: 10.1128/MMBR.00002-17

Zhu, Y.-X., Zhang, Y.-Y., Zhang, X., & Hong, X.-Y. (2023). Antibiotics and temperature alter microbiome assembly and host fecundity in spider mites. *Systematic and Applied Acarology*, *28*(1), 39–52. doi: 10.11158/saa.28.1.5

Paper I

The effect of periodic disturbances and carrying capacity
on the significance of selection and drift
in complex bacterial communities

ISME

www.nature.com/ismecomms

ARTICLE

Check for updates

The effect of periodic disturbances and carrying capacity on the significance of selection and drift in complex bacterial communities

Madeleine S. Gundersen[1✉], Ian Arthur Morelan[2], Tom Andersen[3], Ingrid Bakke[1] and Olav Vadstein[1]

Understanding how periodical disturbances affect the community assembly processes is vital for predicting temporal dynamics in microbial communities. However, the effect of dilutions as disturbances are poorly understood. We used a marine bacterial community to investigate the effect of disturbance (+/−) and carrying capacity (high/low) over 50 days in a dispersal-limited 2 × 2 factorial study in triplicates, with a crossover in the disturbance regime between microcosms halfway in the experiment. We modelled the rate of change in community composition between replicates and used this rate to quantify selection and ecological drift. The disturbed communities increased in Bray–Curtis similarity with 0.011 ± 0.0045 (Period 1) and 0.0092 ± 0.0080 day^{-1} (Period 2), indicating that selection dominated community assembly. The undisturbed communities decreased in similarity at a rate of −0.015 ± 0.0038 day^{-1} in Period 1 and were stable in Period 2 at 0.00050 ± 0.0040 day^{-1}, suggesting drift structured community assembly. Interestingly, carrying capacity had minor effects on community dynamics. This study is the first to show that stochastic effects are suppressed by periodical disturbances resulting in exponential growth periods due to density-independent biomass loss and resource input. The increased contribution of selection as a response to disturbances implies that ecosystem prediction is achievable.

ISME Communications; https://doi.org/10.1038/s43705-021-00058-4

INTRODUCTION

Understanding how ecological assembly processes create temporal patterns in community composition is a major goal in community ecology [1]. After decades of debating whether community assembly follows neutral [2] or niche theory [3], it is now generally accepted that both stochastic and deterministic processes are important for community assembly [1, 4, 5].

Four high-order processes structure community assembly. These are selection, ecological drift, dispersion, and diversification [4, 5]. These four processes have a varying degree of stochasticity and determinism. Selection is deterministic and based on differences in the fitness between populations. This process includes environmental filtering and biological interactions, such as competition and mutualisms. Drift is an entirely stochastic process that arises because there is a non-zero probability that an individual dies before it reproduces [6]. The outcome of drift is a change in the relative abundance of populations and can lead to local extinction if the abundance is low. Dispersion and diversification are two processes that are both stochastic and deterministic. Dispersion refers to an individual's movement from the regional to the local species pool, whereas diversification is the evolution of new strains [4]. The relative contribution of these four processes on community assembly can vary between sites and changes over time [7, 8].

Only experiments with high temporal resolution can evaluate the relative importance of these community assembly processes [9, 10]. During the last decade, studies using high temporal resolution sampling approaches have pointed to stochastic processes as being more important and selection as less important than previously assumed. This observation has been done in habitats such as bioreactors [11, 12], soil [13, 14], and wastewater treatment plants [15]. This increased awareness of stochasticity emphasises the need for more knowledge on temporal variation in the assembly processes.

A primary motivation for studying microbial community assembly is to understand the communities' responses to drivers affecting the high-order assembly processes in order to be able to forecast and manage them [10, 16]. Such control is vital in, for example, treating dysfunctional human gut microbiomes [17], ensuring stability during biological wastewater treatment [18] and providing an optimal microbial environment for fish in aquaculture [19].

Microbial communities often experience disturbances. Disturbances usually involve alterations in the available resources or the biomass concentration in the given environment. To predict the consequence of disturbances on the dynamics of microbial communities' it is essential to understand how the disturbance influences the four assembly processes' relative contributions [7, 10]. Some studies have shown that disturbances affect the relative contribution of the assembly processes [11, 20–25], but conclusions vary depending on the disturbance type and the

[1]Department of Biotechnology and Food Science, NTNU - Norwegian University of Science and Technology, Trondheim, Norway. [2]Independent researcher, Portland, OR, USA. [3]Department of Biosciences, University of Oslo, Oslo, Norway. ✉email: madeleine.gundersen@ntnu.no

Received: 10 June 2021 Revised: 3 September 2021 Accepted: 20 September 2021
Published online: 11 October 2021

ecosystem studied. Zhou et al. formulated two hypotheses describing the community assembly responses to disturbance based on the disturbance type [24]. Disturbances increasing the resource availability are suspected of enhancing the contribution of stochasticity [24] by weakening competition and strengthening priority effects (i.e., the effect of colonisation history) [7, 11, 24, 26]. Their hypothesis regarding biomass-reducing disturbances is contrasting, where deterministic processes should dominate mainly due to increased niche selection [24]. Thus, how disturbances that combine resource increase and biomass loss affect the community assembly is not known.

The maximum biomass an ecosystem can sustain is controlled by the carrying capacity. With regards to community assembly, carrying capacity can affect drift. Lower carrying capacities support lower biomasses, and as drift is density-dependent, more populations are vulnerable to extinction [6]. To our knowledge, no one has investigated how carrying capacity affects community assembly processes.

Microbial microcosms are excellent systems to study the effect of disturbances and carrying capacity on the temporal changes in community assembly. This is due to the short generation time of microorganisms, the potential for high experimental control and the possibility to include many experimental units [27]. In microcosms, one can eliminate dispersal, and if community composition is monitored by clustering 16s-rDNA sequences at a 97% similarity level, speciation is negligible [28]. Consequently, selection and drift are the only assembly processes shaping the bacterial communities [27].

Selection and drift can be quantified by investigating the similarity in community composition between biological replicates in systems without dispersal and speciation (Fig. 1). This approach assumes that if the selection is homogeneous (i.e., there is one stable equilibrium per condition), communities of replicate microcosms should over time become more similar if selection dominates and less similar if drift predominates. Moreover, if selection dominates, one expects the variation in community composition between replicate microcosms to decrease because the communities become more similar over time. Conversely, if drift dominates community assembly, replicates are expected to become less similar, and the variation in compositional similarity will increase with time.

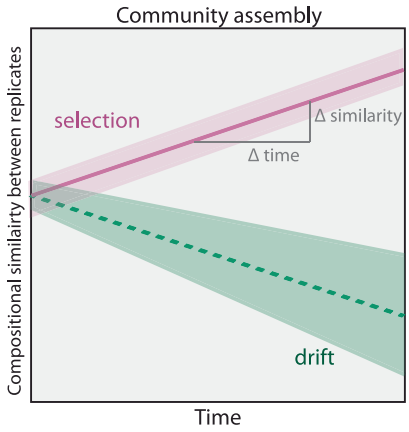


Fig. 1 A conceptual schematic of the temporal changes in community similarity between replicates if drift or selection dominates the community assembly. If selection dominates, the similarity between replicates increases over time, and the variance decrease or be stable. However, if drift dominates, replicates should become less similar over time, and the variance should increase.

In the present study, we aimed at disentangling the effects of disturbance introduced as periodic dilutions (undisturbed versus disturbed microcosms) and carrying capacity (high versus low) on succession and the relative importance of the assembly processes selection and drift. Specifically, we wanted to investigate the following research questions: (1) Do disturbances in the form of dilutions enhance selection or drift? (2) Does lower carrying capacities enhance the contribution of drift? (3) Are the effects of disturbances and carrying capacity dependent on the previous state of the communities?

To answer these questions, we used a 2×2 factorial crossover experimental design with three replicate microbial microcosms for each condition. The marine microbial communities were cultivated for 50 days either in chemostats or with semi-continuous cultivation with a 50-fold dilution every second day. The dilution functioned as a combined disturbance as it both reduced the community size and increased the specific resource supply. We quantified selection and drift using the approach described above, which allowed us to understand the effect of disturbance and carrying capacity on selection and drift.

MATERIALS AND METHODS

Experimental design and sampling scheme

A marine bacterial community was collected from sand-filtered water (~50 µm) collected from 70 m depth in the Trondheimsfjord, Norway (March 2018) and used to inoculate twelve microcosms (500 mL, GLS 80° stirred reactor, Duran, Germany) in a 2×2 factorial crossover design (Fig. 2). Each microcosm contained 250 mL culture that was stirred continuously (MIX 6, 2 mag AG, Germany), supplied with 0.2 µm filtered (Millipore) hydrated air, and kept at 15 °C. The communities were cultivated in f/2 medium [29] with either 0.33 (low carrying capacity, L) or $5 \times 0.33 = 1.67$ mg/L (high carrying capacity, H) of yeast extract, peptone and tryptone. The inorganic nutrients in the f/2 media were 50-fold diluted compared to the original recipe. The medium was either supplied continuously at a dilution rate of 1 day^{-1} (Watson Marlow 520S peristaltic pump) or pulsed by a 1:50 dilution every second day equivalent to a continuous dilution rate of ~2 day^{-1} (Fig. 2a, b). We define the pulsed communities as disturbed (D) and those continuously supplied with medium as undisturbed (U). On day 0, 250 mL of the bacterial community was added to the undisturbed microcosms. In contrast, the communities in the disturbed microcosms were 1:50 diluted in sterile media upon inoculation to a final volume of 250 mL. This disturbance regime was crossed after 28 days so that previously disturbed microcosms were undisturbed the last 22 days (DU) and vice versa (UD). The cultivation regimes are abbreviated as UDH, UDL, DUH, and DUL (Fig. 2c). Each cultivation regime was run in triplicates. The bacterial communities were sampled by filtering ~30 mL of culture through a 0.2 µm filter to a total of 206 samples (2 inoculum and 17 time-points × 4 regimes × 3 replicates) which were stored at −20 °C until further processing. Sampling of the disturbed communities was done right before the dilution.

Extraction of bacterial DNA and 16S-rRNA amplicon sequencing

Bacterial community DNA was extracted with the Qiagen DNeasy PowerSoil DNA extraction kit. The V3-V4 region of the bacterial 16S-rRNA gene was amplified using the broad-coverage primers with Illumina MiSeq adapter sequences Ill338F (5′-TCG-TCG-GCA-GCG-TCA-GAT-GTG-TAT-AAG-AGA-CAG-NNN-NCC-TAC-GGG-WGG-CAG-CAG-3′) and Ill805R (5′-GTC-TCG-TGG-GCT-CGG-AGA-TGT-GTA-TAA-GAG-ACA-GNN-NNG-ACT-ACN-VGG-GTA-TCT-AAK-CC-3′). The reactions were run for 28 cycles (98 °C 15 s, 55 °C 20 s, 72 °C 20 s) with 0.3 µM of each primer, 0.25 mM of each dNTP, 1 mM of $MgCl_2$, 2 µM of 5x Phusion buffer HF, 0.015 units/µL of Phusion Hot Start II DNA polymerase, 1 µL of DNA template and dH_20 to a total volume of 25 µL. The amplicon library was prepared as described previously [30]. In brief, we used the SequalPrep Normalisation plate (96) kit (Invitrogen) to normalise and purify PCR products and the Illumina Nextera XT Index kits (FC-131–2001 and FC-131–2004) for amplicon indexing. The amplicon library was sequenced with V3 reagents by 300 bp paired-end reads on two MiSeq Illumina runs at the Norwegian Sequencing Centre. Illumina sequencing data are deposited at the European Nucleotide Achieve (accession number ERS7182426-ERS7182513).

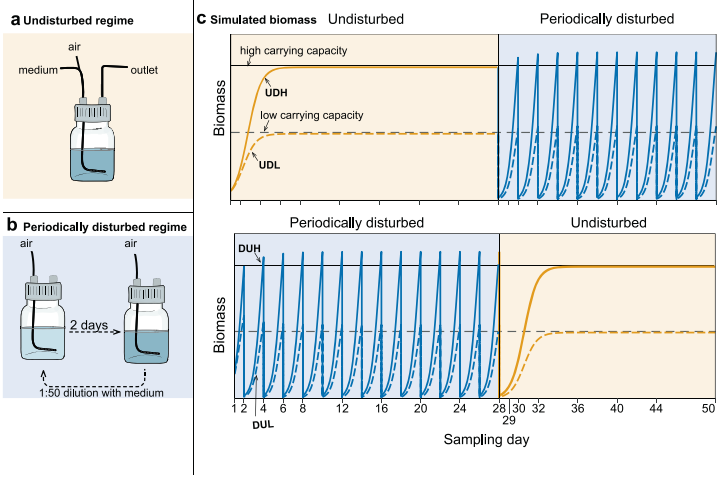


Fig. 2 Schematic overview of the experimental design. a Undisturbed communities (U) received medium continuously at a dilution rate of 1 day^{-1}. **b** Disturbed communities (D) were 50-fold diluted with medium every second day. The dilution acted as a disturbance because the community's biomass was reduced substantially, and resources were introduced as a large pulse. **c** Simulated logistic growth of the microbial communities' biomass given the disturbance regime and carrying capacity (parameters: $\mu = 2.5$ day^{-1}, continuous dilution = 1 day^{-1} or semi-continuous 1:50 dilution every 2nd day). High carrying capacity is indicated as a solid black line, whereas low carrying capacity is presented as a dashed black line. When the communities are undisturbed, the biomass is expected to be at or near carrying capacity. The disturbance by dilution lowered the community's biomass by a factor of 50, bringing the community considerably below the carrying capacity, resulting in close to exponential growth between dilutions. Triplicate microcosms were operated over 50 days for each experimental condition, and the disturbance regime was switched after 28 days. The groups are abbreviated as UDH, UDL, DUH, and DUL, in which the first letter indicates the disturbance regime in period 1 (day 1–28), the second the disturbance regime in period 2 (day 29–50) and the third letter the carrying capacity of the media (high or low).

Processing of sequence data

We used the USEARCH pipeline (v11) to process the Illumina sequence data [31]. Briefly, using the command fastq_mergepairs, paired ends were merged, and primer sequences and reads shorter than 400 bp were removed. The data was quality filtered using the command fastq_filter with an expected error parameter of 1, and singletons were removed. We used the UPARSE-OTU algorithm to remove chimaeras and cluster OTUs at the 97% similarity level [32]. Taxonomy was assigned to the OTUs using the Sintax command with the RDP reference dataset (RPD training data set version16) at an 80% confidence threshold [33, 34].

Analysis of diversity and differential abundance testing

The resulting OTU-table was further analysed in R (version 3.6.1) [35]. All R-code is provided at https://github.com/madeleine-gundersen/disturcance-cc-assembly. We first evaluated the sequencing effort with the function rarecurve() in the vegan package (version 2.5–6) [36]. Then the data were normalised by averaging 1000 rarefied datasets created by randomly sub-sampling 10,000 reads without replacement using phyloseq_mult_raref() from the package metagMisc (version 0.0.4) (https://github.com/vmikk/metagMisc/).

Alpha diversity was estimated as Hill diversity of order 0–2 [37] with the function renyi() in vegan. Bray–Curtis and Sørensen similarity indices were used to quantify beta diversity [38]. The variance in beta-diversity was ordinated with Principal Coordinate analysis (PCoA) [39]. Permutational multivariate analysis of variance (PERMANOVA) was used to test if sample groups significantly differed in community composition. The effect size of variables was evaluated with the R^2-value estimated with PERMANOVA [40]. To determine which OTUs increased in abundance due to the disturbance regimes, we performed a differential abundance test with DeSeq2 [41]. We used the non-normalised OTU-table as input to the DeSeq2 analysis. Only samples from the last 2 weeks of the cultivation periods were included in the analysis as PCoA ordinations indicated that the communities had stabilised. First, the abundance data were normalised using the median ratio method. DeSeq2 was then run with the Wald significance test assuming the negative binomial distribution. All *p*-values were FDR corrected.

Estimation of selection and drift on community composition

We developed a new approach to quantify the contribution of selection and drift during succession in highly controlled experimental settings where dispersal and speciation can be negligible (Fig. 1). Our approach is based on a three-step analytical process. First, the similarity in community composition between replicate pairs is calculated at each sampling day. Then the change in similarity is regressed again time. Finally, the slope of the temporal change in similarity is used to quantify selection and drift. Selection will result in communities becoming more similar with time, resulting in a positive or neutral slope. In contrast, drift causes communities to become less similar over time, manifested as negative slopes. In addition to the slope, the variation in similarity measurements can strengthen the conclusions as selection should decrease variation. In contrast, drift should increase the variation.

We calculated pair-wise community similarities between replicate microcosms at each sampling day, using Bray–Curtis and Sørensen similarity indices. In the following, we will use the term "replicate similarity" for this metric. We used a hierarchical Bayesian model approach to estimate the rate of change in the replicate similarity. We chose a Bayesian approach as it has the advantage of accounting for this dataset's hierarchical dependencies, few observations per time point and the observed heteroscedastic variance [26, 42].

We fitted hierarchical linear Bayesian models with replicate similarity as the dependent variable using the brms package (version 2.11.1) [43], which is a user-friendly front-end for the Stan system for Bayesian computing [44]. All models had a random intercept term for the three similarity comparisons (+(1|comparison) in each time and regime combination. We modelled the replicate similarity by a normal distribution with fixed effects on both mean (μ) and standard deviation (σ) using default priors. Fixed effects included 3-way interactions between time, disturbance, and carrying capacity for the mean model, whereas the standard deviation model only had interactions between time and disturbance. We mean-centred the time variable to reduce correlations between fixed effect estimates. MCMC simulations with brms were run on 4 chains with 4000 samples each (2000 for warm-up), giving 8000 posterior samples. To reduce the number of divergent transitions in the MCMC sampling, we increased the value of the adapt_delta parameter to 0.99 (default is 0.95).

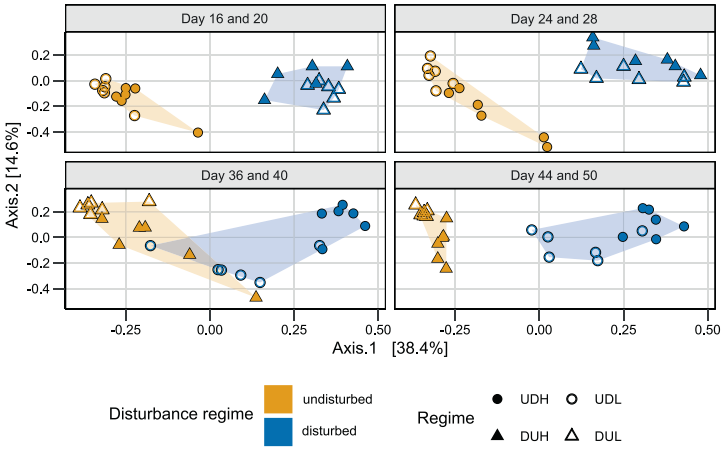


Fig. 3 PCoA ordination based on Bray–Curtis dissimilarity for the bacterial communities at the end of Period 1 (day 16–28) and Period 2 (36–50). The single ordination for these samples was split by sampling-week to highlight the succession based on the disturbance regime. UD (circles) were undisturbed the first 28 days and disturbed the last 22 days, while DU (triangles) were disturbed in the first period and undisturbed in the second. H (filled) and L (empty) indicates high and low carrying capacity, respectively. Colours represent the disturbance regime at sampling, and the shaded area the spread of samples with similar disturbance regimes.

We fitted several models, compared their predictive densities, and selected the model structure with the highest predictability for the temporal development of similarities between replicates. An overview of all estimated models and model selection process is given in the Supplementary material (Supplementary Figs. 1–2, Supplementary Tables 1–3).

We used the package tidybayes (version 2.0.3, http://mjskay.github.io/tidybayes/) to extract posterior samples and the stat_lineribbon() aesthetic from ggplot2 [45] to visualise fixed effect means and credible intervals of model predictions. As explained above, we interpreted community assembly as being dominated by selection if the time effect on the mean of replicate similarity was non-negative (i.e., $\mu\ day^{-1} \geq 0$), and the standard deviation slope was non-positive ($\sigma\ day^{-1} < 0$) (Fig. 1). Conversely, we interpreted a negative slope for the mean and a positive slope for the standard deviation as a community assembly dominated by drift (i.e., $\mu\ day^{-1} < 0$, $\sigma\ day^{-1} > 0$). In cases where the fit met neither of these criteria, we defined the community assembly as a mix of selection and drift.

RESULTS

To study the effect of the periodical disturbance and carrying capacity on community succession and the assembly processes, we cultured marine microbial communities under the DUH, DUL, UDH, and UDL cultivation regimes and characterised their temporal dynamics using 16S-rDNA amplicon sequencing. The dataset contained a total of 12,945,783 sequence reads with a mean of 63,460 reads (± 31,411 SD) per sample. The dataset was normalised to 10 000 reads per sample (Supplementary Fig. 3). The Hill alpha diversity of order 0, 1, and 2 of the normalised dataset correlated well with the non-normalised dataset ($p < 0.05$). The slopes of linear regressions between the alpha diversities of these datasets were close to one, indicating that the normalised-emulated the non-normalised dataset (Supplementary Fig. 4). The samples from the first sampling day were removed from the dataset because the richness dropped 43% from day 1 to 2 (Supplementary Fig. 5). This reduction was probably an adaption of the original seawater community to the culture conditions. During the rest of the experiment, the richness was relatively stable, and a total of 739 OTUs were observed for the normalised OTU table (Supplementary Fig. 5).

The disturbance regime drove succession

The community succession differed between the cultivation regimes, as indicated by PCoA ordinations based on both Bray–Curtis (Fig. 3: day 16–28, 36–50, Supplementary Fig. 6: day 2–50) and Sørensen dissimilarities (Supplementary Fig. 7: day 16–28, 36–50). Disturbance accounted for over 44 and 50% of the variation in Bray–Curtis dissimilarities at the end of Period 1 and 2, respectively (R^2-effect size, $p < 0.001$, PERMANOVA). Carrying capacity accounted for only 6 ($p = 0.16$) and 11% of the variation ($p = 0.04$) for the two periods. A fascinating observation was that switching the disturbance regime reversed the community succession from the undisturbed ordination space to the disturbed one and vice versa (Fig. 3, Supplementary Fig. 8). These ordinations indicated that disturbance was the main contributor to the succession and that carrying capacity had less effect.

Comparing the replicate similarity at the start and the end of each cultivation period indicated that the communities became more similar during disturbance than when undisturbed (Fig. 4a). For the disturbed communities, the Bray–Curtis similarity increased by 138% during Period 1 (DU) and 46% during Period 2 (UD). In contrast, the undisturbed communities decreased in similarity by 47% during Period 1 (UD) and increased by only 3.9% during Period 2 (DU, Fig. 4a). We investigated the replicate similarity change over time to determine whether selection or drift structured these successional patterns.

Selection dominated during disturbance

We used a Bayesian hierarchical model approach to estimate the replicate similarity change over time, and based on this, we examined whether selection or drift dominated the successions. The deterministic process selection should result in communities increasing in replicate similarity over time. This increased similarity will also result in a decrease in variation between similarity measurements. In contrast, the random process drift would reduce the replicate similarity and increase the variation over time (see Fig. 1 and Materials and method for more information).

For the disturbed communities in Period 1 (DU), the posterior-distributions of the model parameters revealed that the replicate similarity increased with $0.011 \pm 0.0045\ day^{-1}$ (mean ± SD), whereas the standard deviation decreased $0.054 \pm 0.010\ day^{-1}$ (Fig. 4b, c). This increased replicate similarity and decreased standard deviation over time indicate that selection was the dominating assembly process (Fig. 5a). Moreover, we observed the same trends for the disturbed communities in Period 2 (UD) with

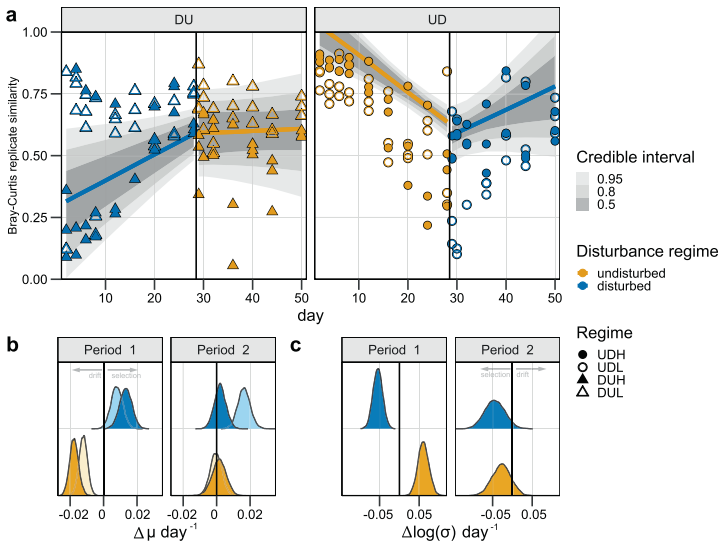


Fig. 4 The Bray–Curtis based models and coefficient estimates for the change in the similarity between replicates over time. a The similarity between replicate communities as a function of time. The models for the replicate similarity change over time are presented as lines with the 0.5, 0.8, and 0.95 credible intervals around it. The observed data used as the response variable in the models are presented as points. UD (circles) were undisturbed the first 28 days and disturbed the last 22 days, while DU (triangles) were disturbed in the first period and undisturbed in the second. H (filled) and L (empty) indicates high and low carrying capacity, respectively. Colours represent the disturbance regime at sampling. **b** The posterior distributions of the expected replicate similarity (μ) change per day given the interaction between time, the disturbance regime and carrying capacity. The distribution reflects all 8000 estimated replicate similarity changes that would give the observed data. Light and dark colours indicate low and high carrying capacity, respectively. The colour indicates the disturbance regime at sampling. **c** The posterior distributions for the change in standard deviation per day given the interaction between time and disturbance regime. The distribution reflects all 8000 estimated standard deviation changes per day that would give the observed data. The colour indicates the disturbance regime at sampling.

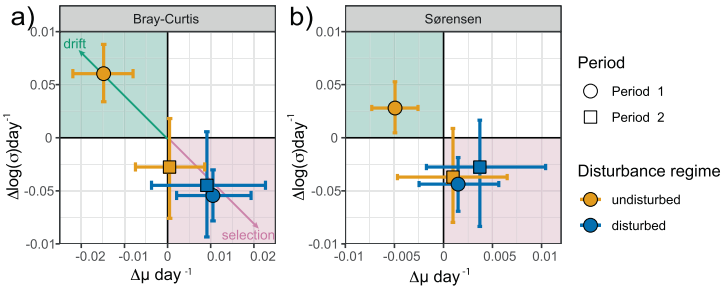


Fig. 5 Model estimates for the replicate similarity and change in standard deviation. The mean change in replicate similarity (μ) over time and mean change in standard deviation (σ) (points) and the accompanying 95% credible intervals for each estimate, as inferred using the Bayesian hierarchical model approach, based on **a** Bray–Curtis and **b** Sørensen based models. The green area indicates the coordinate space where drift dominates, while the pink areas indicate where selection is dominating. Point colour indicates the disturbance regime and shapes the cultivation period.

an increase in replicate similarity of 0.0092 ± 0.0080 day^{-1} and a decrease in the standard deviation of -0.044 ± 0.025 day^{-1} (Fig. 4b, c). This coherent observation strengthens the conclusion that selection dominated community assembly during disturbances.

The modelling results were different for the undisturbed communities. During Period 1 (UD), the replicate similarity rate decreased by -0.015 ± 0.0038 day^{-1} and had a temporal increase in variation of 0.061 ± 0.014 day^{-1} (Fig. 4b, c), indicating that drift dominated the community assembly (Fig. 5a). For the communities that switched from a disturbed to an undisturbed regime (DU) in Period 2 the dominating assembly process was less obvious. The replicate similarity rate was relatively stable with a mean of 0.00050 ± 0.00400 day^{-1} and a decrease in the standard deviation of -0.028 ± 0.024 (Fig. 4b, c). These values categorise the assembly as selection (Fig. 5). However, comparing the replicate similarity rate of the communities from Period 1 to the one in Period 2 shows that the rate decreased substantially. The average similarity rate transitioned from the selection-coordinate space towards the one where drift dominates (Fig. 5a).

The results were similar for models based on the Sørensen similarity, with an overall increase in replicate similarity over time

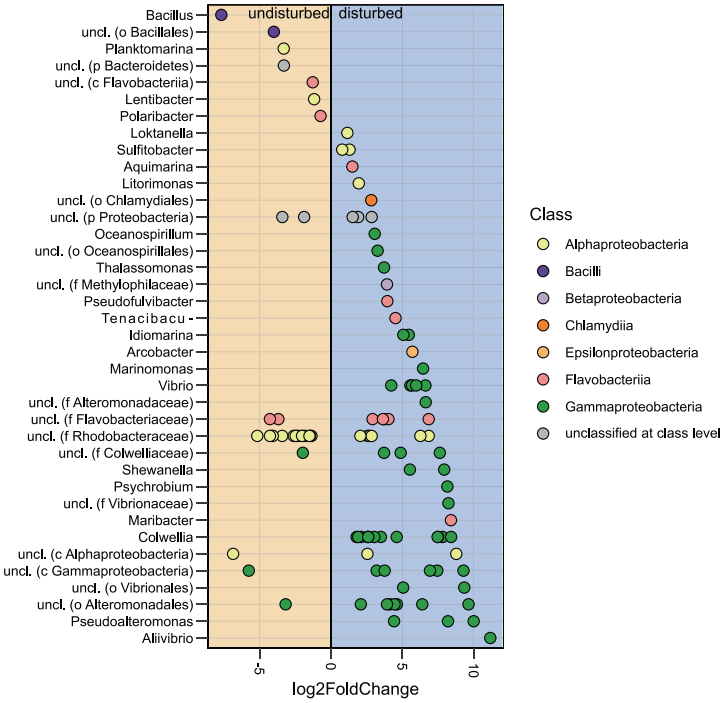


Fig. 6 The log2 fold change in relative abundance between the disturbed and undisturbed communities during the last 2 weeks of the cultivation periods (week 3–4 and 6–7). Only OTUs with a significance level lower than 0.05 are shown (FDR-adjusted DESeq2 *p*-values). Each point represents an OTU coloured by the class classification. OTUs were grouped according to the genus level. The lowest taxonomic classification obtained is indicated in parenthesis for OTUs that could not be classified at the genus level (p phylum, c class, o order, f family, g genus). OTUs with higher abundance during disturbance are in the blue shaded area, whereas those with higher abundance when the environment was undisturbed are in the orange shaded area.

for the disturbed regimes of 0.0015 ± 0.0021 day^{-1} in Period 1 (DU) and 0.0041 ± 0.0040 day^{-1}in Period 2 (UD), and a decrease in the standard deviation of the replicate similarity over time (Fig. 5b, Supplementary Fig. 9). For the undisturbed communities, there was a slight temporal decrease in Sørensen replicate similarity at a rate of -0.0049 ± 0.0012 day^{-1} in Period 1 (UD), whereas in Period 2 there was an insignificant change in replicate similarity (0.00096 ± 0.0029 day^{-1}; DU). These results supported the findings based on the Bray–Curtis similarity; drift dominated assembly for the undisturbed communities, whereas selection dominated when the communities were disturbed.

Gammaproteobacteria increased in relative abundance during disturbance

The PCoA ordination and the replicate similarity models showed that the disturbance regime impacted the assembly processes. We performed a DeSeq2 differential analysis to elucidate which OTUs had significantly different abundances between the disturbed and undisturbed regime. This analysis revealed that 107 of the 535 OTUs contributed significantly ($p < 0.05$) to differences in community composition between the disturbed and the undisturbed regimes. These OTUs were grouped at the genus level (Fig. 6). Interestingly, around 60% of these genera included only one OTU. For genera with more OTUs affected, the general trend was that the OTUs responded similarly to the disturbance regime (i.e. either positive or negative fold change in relative abundance). For example, all 13 OTUs classified as *Colwellia* and all 5 OTUs classified as *Vibrio* had higher abundance during disturbance. However, this was not the case for all the groups. For example, of the 21 OTUs classified to *Rhodobacteraceae*, 6 were in higher abundances during the disturbed periods, whereas 15 were more abundant during undisturbed periods. Thus, some genera's OTU abundances appeared to respond to the disturbance regime coherently, whereas others did not. Of the 107 OTUs significantly affected by the disturbance regime, 72% had increased abundances when the environment was disturbed. Especially noteworthy was the *Gammaproteobacteria*, where 94% of the OTUs significantly affected by the disturbance regime had higher abundances during disturbance with up to an 11.2 fold-change.

DISCUSSION

Predicting community responses to ecosystem changes is essential for improving ecosystem management. From an industrial perspective, we are dependent on stable microbial communities that perform well. Moreover, we live in a time where humans create disturbances at various levels in natural ecosystems. It is therefore important to comprehend the consequences of our activity. To predict the community response to external forces, we need to understand how different ecosystems affect the community assembly processes.

We aimed to fill the knowledge gap on how carrying capacity and periodical disturbances affect the community assembly. It has previously been shown that the carrying capacity affects the community composition [46]. However, its effect on the assembly processes has remained unclear. Ecosystems with a lower carrying capacity support lower community size. Because the outcome of

drift is density-dependent [6], communities with a low carrying capacity should have more populations vulnerable to drifting to extinction. However, our five-times difference in carrying capacity between cultivation regimes did not result in apparent differences in community assembly. The only exception was for the disturbed communities in Period 2, where the low carrying capacity regime (UDL) indicated a stronger influence of selection than the high (UDH; Fig. 4b). This observation was surprising as we hypothesised that drift might be more pronounced in systems with lower carrying capacity. In conclusion, the minor effects of carrying capacity observed for the replicate similarity rate for the undisturbed communities suggest that the effect of carrying capacity should be investigated further, including larger differences in carrying capacity.

The effect of the disturbance regime on the microbial community assembly was more evident. The disturbance we investigated was a substantial dilution of the microcosm's inoculum. The dilution has two significant effects: the community size is reduced, and the concentration of resources increases strongly for the remaining individuals. These two changes are relevant in natural and human-created ecosystems, where resource supply vary due to natural processes (e.g. patchiness and floods) and human activity (e.g. eutrophication and saprobiation).

Investigating the temporal community composition through ordinations can reveal overall successional trajectories [47]. We found that whereas the PCoA ordinations indicated an overall deterministic trajectory for the undisturbed communities, the replicate similarity rate indicated that drift dominated the community assembly. This was evident for the microcosms starting with undisturbed culture conditions (UD $\Delta\mu > 0$; Fig. 5). However, the results were less evident for the communities going from disturbed to undisturbed conditions (DU) as the replicate similarity rate was around zero. Nonetheless, there was an apparent decrease in the replicate similarity rate when going from disturbed ($\Delta\mu$ 1.1×10^{-2}) to undisturbed conditions ($\Delta\mu$ 5×10^{-4}).

The strength and unique feature of our experiment is the crossed design of the disturbance regimes. This crossing considerably increases the robustness of the conclusions drawn from the data. First, during the first period, all microcosms were inoculated with the same community, but in the second period, the twelve communities had assembled individually for 28 days. We could therefore investigate the effects of our experimental variables on drift and selection with different starting conditions. The temporal trends in the data were found to be independent of the starting condition, substantially increasing the strength of our conclusion.

Second, subjecting the communities to the opposite disturbance regime in Period 2 supports that we had stable attractors in our systems. An attractor is a point or a trajectory in the state space of a dynamical system. If the attractor is locally stable, the system will tend to evolve toward it from a wide range of starting conditions and stay close to it even if slightly disturbed [48]. We observed locally stable attractors based on the disturbance regime and thus one stationary phase for each disturbance regime. Some ecological systems show dramatic regime shifts between alternative stationary states in response to changes in an external driver [49]. Such systems typically exhibit hysteresis in the sense that they will not return directly to the original state by an opposite change in the driver. We found that community composition was reversible and dependent on the disturbance regime, as highlighted by the Bray–Curtis ordinations (Fig. 4). This reversibility indicates that the community changes we observed were not catastrophic bifurcations or regime shifts and that it is unlikely that the systems contain multiple stationary states within the same disturbance regime. We think this gives strong support for assuming that drift is the main driver for divergence in the community composition and that selection towards alternative attractors probably plays a minor role. Thus, we can conclude that shifting from a disturbed to an undisturbed ecosystem increased the contribution of drift. Our observations corroborate other investigations of bioreactors [15, 50] and simulations [51] that report that stochasticity is fundamental for the assembly of communities. However, the finding that drift was important for structuring the undisturbed microcosms was unexpected.

In dispersal-limited communities where resources are supplied continuously, such as in the undisturbed communities examined here, the selective process competition has been hypothesised to be high [7]. However, our experimental environment offered little variation in the resources provided, as the medium provided was the same throughout the experiment. This may have led to populations becoming "ecologically equivalent", meaning that their fitness difference was too small to result in competitive exclusion on the time scale of our experiment [5, 52]. Under these assumptions, community assembly is similar to the neutral model in which the growth rates of the community members are comparable [53].

During disturbances, we found that selection dominated community assembly. Our results support Zhou et al. hypothesis stating that determinism should increase due to biomass loss in dispersal-limited communities [24]. However, they oppose their other hypothesis stating that nutrient inputs should increase stochasticity [24], making low abundant populations vulnerable to local extinction [6, 7]. During the disturbances, the Sørensen similarity between replicates was stable or increasing, indicating that the periodical disturbance did not result in the extinction of low abundant populations. Instead, it appears that the dilution removed competition for some time, resulting in a phase where all populations got "a piece of the cake". Several studies have observed increased stochasticity as a result of increased resource availability [7, 11, 24, 26]. However, we found that disturbances resulting in periods with exponential growth due to density-independent loss of individuals and high resource input suppressed the effect of stochastic processes. This exponential growth period without competition would enable more populations to stay above the detection limits of the 16S-rDNA-sequencing method.

More OTUs were enriched under the disturbed regime than under the undisturbed. During the disturbance, the microcosms were diluted ~2 day^{-1}, whereas the dilution factor was 1 day^{-1} during the undisturbed regime. We cannot assume steady-state in the disturbed microcosms, but it was interesting to see a substantial increase in the abundance of OTUs classified as *Gammaproteobacteria*. *Gammaproteobacteria* include many opportunists [54] that appeared to exploit the resource surplus following the disturbance. This opportunistic lifestyle fits within the r- and K-strategist framework [55].

Organisms with high maximum growth rates but low competitive abilities are classified as r-strategists. These r-strategists are superior in environments where the biomass is below the carrying capacity. On the other hand, K-strategists are successful in competitive environments due to their high substrate affinity and resource specialisation [56]. Based on the taxonomic responses, it appears as disturbances in the form of dilutions selected for r-strategists, whereas the undisturbed regime selected for K-strategists. The r-strategists selected for during the disturbance periods included genera such as *Vibrio* and *Colwellia* [57], and the genus *Vibrio* includes many pathogenic strains [58]. Thus, our findings may have implications for land-based aquaculture systems where conditions favouring r-strategists is linked to high mortality and reduced viability of fish [56].

The DeSeq2 results pose some new questions regarding the link between phylogeny and niche fitness. Generally, ecologists assume that closely related taxa have similar niches, as they have a common evolutionary history and, thus, similar physiology

[59, 60]. For example, here, OTUs belonging to *Gammaproteobacteria* co-occurred when the environment was disturbed. However, for other classes such as *Alphaproteobacteria* and *Flavobacteria*, the OTUs responded differently to the disturbance regimes, despite belonging to the same class. This lack of phylogenetically coherent response indicates that the paradigm of correlation between phylogeny and niche requires further studies.

This study was performed on complex marine microbial communities cultivated under controlled experimental conditions. We found that undisturbed environments enhanced the contribution of drift on community assembly and that disturbances increased the effect of selection. These observations might be different in more diverse ecosystems such as soils or the human gut. In such ecosystems, the microbes are more closely associated with, for example, soil particles or attached to the gut lining. It has been shown that the biofilm-associated and planktonic microbial communities have different community compositions [61]. Consequently, the community assembly processes may be affected differently by environmental fluctuations. Our experimental variables should therefore be tested in other ecosystem settings to verify our conclusions.

To our knowledge, this study is the first to experimentally estimate the effect of periodical disturbances and carrying capacity on community assembly in dispersal-limited ecosystems. We observed that carrying capacity had little effect on community assembly and that undisturbed communities were structured more by drift than disturbed systems dominated by selection. Using an experimental crossover design for the disturbance regime, we showed that these observations were independent of the initial community composition. Our experiment illustrates that cultivating complex natural microbial communities under lab conditions allowed us to test ecologically relevant system variables and draw robust conclusions.

REFERENCES

1. Zhou J, Ning D. Stochastic community assembly: does it matter in microbial ecology? Microb Mol Biol Rev. 2017;81:e00002–17.
2. Hubbell SP. The unified neutral theory of biodiversity and biogeography. Princeton New Jersey: Princeton University Press; 2001.
3. Chesson P. Mechanisms of maintenance of species diversity. Ann Rev Ecol Syst. 2000;31:343–66.
4. Vellend M. Conceptual synthesis in community ecology. Quarterly Rev Biol. 2010;85:183–206.
5. Nemergut DR, et al. Patterns and processes of microbial community assembly. Microbiol Mol Biol Rev. 2013;77:342–56.
6. Vellend M. The theory of ecological communities. New Jersey: Princeton University Press, Princeton; 2016. p. 57.
7. Dini-Andreote F, Stegen JC, van Elsas JD, Salles JF. Disentangling mechanisms that mediate the balance between stochastic and deterministic processes in microbial succession. Proc Natl Acad Sci. 2015;112:E1326 LP–E1332.
8. Langenheder S, Berga M, Östman Ö, Székely AJ. Temporal variation of β-diversity and assembly mechanisms in a bacterial metacommunity. ISME J. 2012;6:1107–14.
9. Gonzalez A, et al. Characterizing microbial communities through space and time. Curr Opin Biotechnol. 2012;23:431–6.
10. De Vrieze J, et al. Stochasticity in microbiology: managing unpredictability to reach the Sustainable Development Goals. Microbial Biotechnol. 2020;13:829–43.
11. Zhou J, et al. Stochastic Assembly Leads to Alternative Communities with Distinct Functions in a Bioreactor Microbial Community. mBio. 2013;4:e00584–12.
12. Liébana R, et al. Combined Deterministic and Stochastic Processes Control Microbial Succession in Replicate Granular Biofilm Reactors. Environ Sci Technol. 2019;53:4912–21.
13. Rigonato J, et al. Temporal assessment of microbial communities in soils of two contrasting mangroves. Brazilian J Microbiol. 2018;49:87–96.
14. Guo X, et al. Climate warming leads to divergent succession of grassland microbial communities. Nat Climate Change. 2018;8:813–8. *2018 8:9*
15. Ofiţeru ID, et al. Combined niche and neutral effects in a microbial wastewater treatment community. PNAS. 2010;107:15345–50.
16. Rittmann BE. Microbial ecology to manage processes in environmental biotechnology. Trends Biotechnol. 2006;24:261–6.
17. Shafquat A, Joice R, Simmons SL, Huttenhower C. Functional and phylogenetic assembly of microbial communities in the human microbiome. Trends Microbiol. 2014;22:261–6.
18. Graham DW, Smith VH. Designed ecosystem services: application of ecological principles in wastewater treatment engineering. Designed Ecosyst Serv. 2004;2:199–206.
19. Vadstein O, et al. Managing the microbial community of marine fish larvae: a holistic perspective for larviculture. Front Microbiol. 2018;9:1820.
20. Chase JM. Drought mediates the importance of stochastic community assembly. PNAS. 2007;104:17430 LP–17434.
21. Herold M, et al. Integration of time-series meta-omics data reveals how microbial ecosystems respond to disturbance. Nat Commun. 2020;11:1–14.
22. Vass M, Székely AJ, Lindström ES, Langenheder S. Using null models to compare bacterial and microeukaryotic metacommunity assembly under shifting environmental conditions. Scientific Reports. 2020;10:1–13.
23. Fillinger L, Hug K, Griebler C. Aquifer recharge viewed through the lens of microbial community ecology: Initial disturbance response, and impacts of species sorting versus mass effects on microbial community assembly in groundwater during riverbank filtration. Water Res. 2021;189:116631.
24. Zhou J, et al. Stochasticity, succession, and environmental perturbations in a fluidic ecosystem. PNAS. 2014;111:E836–45.
25. Ferrenberg S, et al. Changes in assembly processes in soil bacterial communities following a wildfire disturbance. ISME J. 2013;7:1102–11.
26. Shade A et al. Fundamentals of microbial community resistance and resilience. Front Microbiol. 2012;3:417.
27. Jessup CM, et al. Big questions, small worlds: microbial model systems in ecology. Trends Ecol Evol. 2004;19:189–97.
28. Burns AR, et al. Contribution of neutral processes to the assembly of gut microbial communities in the zebrafish over host development. ISME J. 2015;10:655–64.
29. Smith WL, Chanley MH Culture of marine invertebrate animals: proceedings of the conference on culture of marine invertebrate animals. New York: Plenum Press; 1975.
30. Vestrum RI, Attramadal KJK, Vadstein O, Gundersen MS, Bakke I. Bacterial community assembly in Atlantic cod larvae (Gadus morhua): contributions of ecological processes and metacommunity structure. FEMS Microbiol Ecol. 2020;96:1–12.
31. Edgar RC. Search and clustering orders of magnitude faster than BLAST. Bioinformatics. 2010;26:2460–1.
32. Edgar RC. UPARSE: highly accurate OTU sequences from microbial amplicon reads. Nat Methods. 2013;10:996–8.
33. Cole JR, et al. Ribosomal Database Project: data and tools for high throughput rRNA analysis. Nucleic Acids Res. 2013;42:D633–42.
34. Edgar RC. SINTAX: a simple non-Bayesian taxonomy classifier for 16S and ITS sequences. bioRxiv. 2016;74161.
35. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2020. https://www.R-project.org/.
36. Dixon P. VEGAN, a package of R functions for community ecology. J Vegetation Sci. 2003;14:927–30.
37. Lucas R et al. A critical evaluation of ecological indices for the comparative analysis of microbial communities based on molecular datasets. FEMS Microbiol Ecol. 2017;93:1–15.
38. Legendre P, Cáceres Mde. Beta diversity as the variance of community data: dissimilarity coefficients and partitioning. Ecol Lett. 2013;16:951–63.
39. Ramette A. Multivariate analyses in microbial ecology. FEMS Microbiol Ecol. 2007;62:142.
40. Anderson MJ. A new method for non-parametric multivariate analysis of variance. Austral Ecol. 2001;26:32–46.
41. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15:1–21.
42. Carpenter SR, Brock WA. Rising variance: a leading indicator of ecological transition. Ecol Lett. 2006;9:311–8.
43. Bürkner P-C. brms: An R Package for Bayesian Multilevel Models Using Stan. J Stat Softw. 2017;80:1–28.
44. Stan Development Team. YEAR. Stan Modeling Language Users Guide and Reference Manual, VERSION. https://mc-stan.org.
45. Wickham H. ggplot2: Elegant Graphics for Data Analysis. New York: Springer-Verlag; 2016. vol 2. p. 1–266.
46. Attramadal KJK, et al. Microbial maturation of intake water at different carrying capacities affects microbial control in rearing tanks for marine fish larvae. Aquaculture. 2016;457:68–72.

47. Fukami T, Bezemer TM, Mortimer SR, Putten WHvander. Species divergence and trait convergence in experimental plant community assembly. Ecol Lett. 2005;8:1283–90.
48. Attractor - from Wolfram MathWorld. https://mathworld.wolfram.com/Attractor.html. Accessed Jun 2021.
49. Scheffer M, Carpenter S, Foley JA, Folke C, Walker B. Catastrophic shifts in ecosystems. Nature. 2001;413:591–6.
50. Yuan H, Mei R, Liao J, Liu W-T. Nexus of Stochastic and Deterministic Processes on Microbial Community Assembly in Biological Systems. Front Microbiol. 2019;10:1536.
51. Evans S, Martiny JBH, Allison SD. Effects of dispersal and selection on stochastic assembly in microbial communities. ISME J. 2017;11:176–85.
52. Leibold MA, Mcpeek MA. Coexistence Of The Niche And Neutral Perspectives In Community Ecology. Ecology. 2006;87:1399–410.
53. Sloan WT, et al. Quantifying the roles of immigration and chance in shaping prokaryote community structure. Environ Microbiol. 2006;8:732–40.
54. Fuhrman JA, Hagström Å. Bacterial and Archaeal Community Structure and its Patterns. Microbial Ecology of the Oceans: Second Edition. 2008; 45–90.
55. Andrews J, Harris R. r- and K-Selection and Microbial Ecology. Adv Microbial Ecol. 1986;9:99–147.
56. Vadstein O, Attramadal KJK, Bakke I, Olsen Y. K-Selection as Microbial Community Management Strategy: A Method for Improved Viability of Larvae in Aquaculture. Front Microbiol. 2018;9:2730.
57. Liu S, Wawrik B, Liu Z. Different Bacterial Communities Involved in Peptide Decomposition between Normoxic and Hypoxic Coastal Waters. Front Microbiol. 2017;0:353.
58. Baker-Austin C, et al. Vibrio spp. infections. Nat Rev Disease Primers. 2018;4:1–19.
59. Coutinho FH, et al. Niche distribution and influence of environmental parameters in marine microbial communities: a systematic review. PeerJ. 2015;3:e1008.
60. Tromas N, et al. Niche Separation Increases With Genetic Distance Among Bloom-Forming Cyanobacteria. Front Microbiol. 2018;0:438.
61. Miao L, et al. Distinct community structure and microbial functions of biofilms colonizing microplastics. Sci Total Environ. 2019;650:2395–402.

ACKNOWLEDGEMENTS

This study was part of the ERA-Net COFASP project "MicStaTech", funded by the Research Council of Norway (Contract 247558). Financial support was also provided by NTNU, Faculty of Natural Sciences, as a PhD scholarship to MSG and a Fulbright scholarship to IAM. We would like to thank T. Frede Thingstad and the Brendan Bohannan group for valuable comments on an earlier draft of this manuscript.

AUTHOR CONTRIBUTIONS

IM, IB, and OV designed the study. MG and IM conducted the study. MG analysed the data with input from TA and OV. MG prepared the initial draft of the manuscript before all authors contributed to writing and editing the final manuscript.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s43705-021-00058-4.

Correspondence and requests for materials should be addressed to Madeleine S. Gundersen.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Paper II

Bacterial community assembly in Atlantic cod larvae (*Gadus morhua*): contributions of ecological processes and metacommunity structure

FEMS Microbiology Ecology, 96, 2020, fiaa163

doi: 10.1093/femsec/fiaa163
Advance Access Publication Date: 20 August 2020
Research Article

RESEARCH ARTICLE

Bacterial community assembly in Atlantic cod larvae (*Gadus morhua*): contributions of ecological processes and metacommunity structure

Ragnhild I. Vestrum†, Kari J. K. Attramadal, Olav Vadstein, Madeleine Stenshorne Gundersen and Ingrid Bakke*

Department of Biotechnology and Food Science, NTNU - Norwegian University of Science and Technology, Trondheim, Norway

*Corresponding author: NTNU Department of Biotechnology and Food Science, 7491 Troneheim, Norway. E-mail: Ingrid.bakke@ntnu.no

One sentence summary: Bacterial community assembly in Atlantic cod larvae is driven by deterministic and stochastic processes and influenced by the bacterial metacommunity structure.

Editor: Julie Olson
†Ragnhild I. Vestrum, http://orcid.org/0000-0001-9012-4709

ABSTRACT

Many studies demonstrate the importance of the commensal microbiomes to animal health and development. However, the initial community assembly process is poorly understood. It is unclear to what extent the hosts select for their commensal microbiota, whether stochastic processes contribute, and how environmental conditions affect the community assembly. We investigated community assembly in Atlantic cod larvae exposed to distinct microbial metacommunities. We aimed to quantify ecological processes influencing community assembly in cod larvae and to elucidate the complex relationship between the bacteria of the environment and the fish. Selection within the fish was the major determinant for community assembly, but drift resulted in inter-individual variation. The environmental bacterial communities were highly dissimilar from those associated with the fish. Still, differences in the environmental bacterial communities strongly influenced the fish communities. The most striking difference was an excessive dominance of a single OTU (*Arcobacter*) for larvae reared in two of the three systems. These larvae were exposed to environments with higher fractions of opportunistic bacteria, and we hypothesise that detrimental host–microbe interactions might have made the fish susceptible to *Arcobacter* colonisation. Despite strong selection within the host, this points to a possibility to steer the metacommunity towards mutualistic host–microbe interactions and improved fish health and survival.

Keywords: microbiota; microbial ecology; community assembly; Atlantic cod; metacommunity; ecological processes

INTRODUCTION

There has been a great advancement in our understanding of the microbiota associated with animal hosts and its roles in host health and development. The gut microbiota plays important roles in epithelial differentiation and maturation (Naito *et al.* 2017), contributes to metabolism of nutrients and xenobiotics (Semova *et al.* 2012; Sonnenburg and Bäckhed 2016; Koppel, Maini Rekdal and Balskus 2017) and is essential for the development of the immune system (Hiippala *et al.* 2018). The indigenous microbiota also protects the host by preventing colonisation by harmful bacteria (Lazado *et al.* 2011; Lazado and Caipang

Received: 4 March 2020; **Accepted:** 12 August 2020

2014; Hiippala *et al.* 2018). Dysbiosis in the gut microbiota is associated with an increasing number of diseases (Rogers *et al.* 2016; Brugman *et al.* 2018), and the microbiota can affect both growth and survival (Vadstein *et al.* 2018). Consequently, there is great interest in understanding factors and processes that determine the composition of the animal gut microbiota, which have been proposed to include host genetics, developmental stage (Bonder *et al.* 2016), diet (David *et al.* 2013), environmental microbes (Fujimura *et al.* 2014) and selection in the host (Rawls *et al.* 2006). However, in natural habitats these factors are often interacting, and are thus hard to study. For example, it was only recently revealed that host genetics has a relatively small impact on the gut microbiota of humans compared to the impact of environmental factors such as geographical location, diet and age (Jackson *et al.* 2018; Rothschild *et al.* 2018).

Fish larvae are well suited for experimental studies of community assembly and dynamics of vertebrate-associated microbiota, because of their small size, rapid development, possibilities for good sample size and replication, and a wide range of experimental systems and host species (Vestrum *et al.* 2018). During the early colonisation of skin and intestinal system of fish the major source for bacteria entering the fish is assumed to be the surrounding water microbiota (Nayak 2010). The gut is colonised at the mouth opening in young marine fish larvae (Reitan, Natvik and Vadstein 1998). We have previously shown that it is possible to optimise microbial water quality and promote mutualistic host–microbe interactions for cultivated fish by applying ecological theory to set up selection regimes through water treatment (Vadstein *et al.* 2018). Attramadal *et al.* (2014) demonstrated that optimising microbial water quality in both recirculation aquaculture system (RAS) and a microbially matured flow through system (MMS) lead to a 70–90% increase in the survival of Atlantic cod (*Gadus morhua*) larvae compared to traditional flow through rearing systems (FTS). Moreover, we have also shown that cod larval microbiota is affected by the differences in the water microbiota introduced through water treatment (Vestrum *et al.* 2018). Paradoxically, despite this, the fish microbiota is generally highly dissimilar from microbiota in the surrounding water, indicating that selection in the host structures the fish microbiota (Sullam *et al.* 2012; Bakke *et al.* 2015; Giatsis *et al.* 2015). On the other hand, high inter-individual variations between fish in the same environment (Fjellheim *et al.* 2012; Boutin *et al.* 2014) suggest that also stochastic processes like drift and dispersal contribute to microbial community assembly in the host. In general, the relative importance of the various processes and factors influencing the colonisation of animals and especially fish is poorly understood.

Ecological theory has been proposed as a foundation to increase our understanding of host-associated microbiota. Vellend (2016) suggested that four fundamental ecological processes explain patterns in community diversity and composition: selection, dispersal, speciation and drift. Nemergut *et al.* (2013) argue that speciation should be considered diversification in the case of microbial community assembly, as for microbes the species concept is complicated, and the generation of new genetic variation can bring change to a community's dynamics even if new species are not created. Thus, we choose to use diversification when describing this process. Vellend's conceptual synthesis has been found useful also for microbial community assembly (Hanson *et al.* 2012; Nemergut *et al.* 2013), but few have used it for animal hosts. Two studies on zebrafish by Burns and colleagues (Burns *et al.* 2016; Burns *et al.* 2017) concluded that drift and passive dispersal were sufficient to generate substantial variation in the microbiota across individual hosts, and that interhost dispersal can be more important than differences in host immunity. Dispersal of species has the potential to link local communities into what has been defined as a metacommunity (Leibold *et al.* 2004). Traditional metacommunity theory assumes that local communities occur in different patches that are linked through dispersal. Metacommunity theory explains patterns in community composition as a combination of local factors (selection) and regional factors (dispersal between patches). Thus, the patches can also exhibit heterogeneity or similarity over time and space due to variations in dispersal and selection pressure. Miller *et al.* (2018) have proposed extensions to the traditional metacommunity theory to include host–microbiota systems. The authors argue that dispersal occurs both between hosts and between hosts and the environment, and that feedback between the hosts and the environmental microbiota could influence the host microbiota. In addition they propose that the host-associated microbiota may have the ability to change host properties such as fitness and development (Miller, Svanbäck and Bohannan 2018).

In this study, we examine the bacterial community assembly in newly hatched Atlantic cod larvae over a period of 46 days, through detailed characterisation of the bacterial communities of water, feed and individual fish. The fish were reared in triplicate tanks with water from three distinct source bacteria over a period of 30 days, followed by a period of 16 days where all tanks received water with the same bacterial communities. Until 30 days post hatching (dph), we consider each of the three systems, including water, fish and feed, separate metacommunities. We quantified the relative importance of ecological processes under action and elucidated the relationship between the bacterial communities of the water and the fish.

MATERIALS AND METHODS

For this study we analysed the bacteria of cod larvae, water and live feed samples originating from a start feeding experiment with cod larvae, previously described in Attramadal *et al.* (2014). The experimental setup is described briefly below, and further details are given in Attramadal *et al.* (2014). The experiment was carried out at NTNU Sealab within the Norwegian animal welfare act guidelines, in accordance with the Animal Welfare Act of 20 December 1974, amended 19 June 2009, at a facility with permission to conduct experiments on fish (code 93) provided by the Norwegian Animal Research Authority (NARA).

Experimental design and rearing systems

The primary experimental variable in this study was the use of three different water treatments systems to create three distinct microbial communities entering the rearing tanks. These systems were a flow through system (FTS), a microbial maturation system (MMS), and a recirculation aquaculture system (RAS). In FTS and MMS, the carrying capacity (i.e. the maximum cell number that can be sustained over time by the resources available) of the water going into the tanks (incoming water) was significantly lower than in the rearing tank water, while for RAS, it was more or less the same for incoming water and tank water (Attramadal *et al.* 2014; Vadstein *et al.* 2018). Thus, FTS incoming water was considered to represent r-selected microbial communities, and RAS and MMS K-selected microbial communities (Attramadal *et al.* 2014; Vadstein *et al.* 2018). In each system there were three replicate fish rearing tanks (of 160 L) which were maintained from hatching to 30 dph. Thereafter all the nine rearing tanks

received the same microbial water quality (i.e. MMS water) until 46 dph. Details about the water treatments in the three systems are found in Attramadal et al. (2014).

Cod larvae rearing

Fertilised Atlantic cod eggs were received from Nofima marine national breeding station, Havbruksstasjonen i Troms⊘AS. In brief, eggs were disinfected with glutaraldehyde and transferred to the rearing tanks to reach a final density of 100 larvae L^{-1}. The larvae were fed rotifers (*Brachionus* 'Cayman') from day 3–26 dph, *Artemia* nauplii from day 22–32 dph and formulated feed (GEMMA Micro, SKRETTING, Norway) from day 31–46 dph. More details about the egg handling, cod larvae rearing and calculations of the survival of the cod larvae at 32 dph are found in Attramadal et al. (2014) and in Supplementary Table S1, see online supplementary material. The survival of the cod larvae was not calculated at the end of the experiment.

Sampling

Rearing tank water samples (40 mL) from each tank for each system, were collected 4, 8, 17, 30 and 46 dph. In addition, one sample from each system was sampled at 1 dph. Incoming water was sampled on the same days except at 46 dph. Live feed samples (rotifers) from the fish tanks were taken at 8 and 17 dph by collecting 100 mL of tank water, rinsing the feed with sterile water in a sterilized sieve and collecting ∼200 rotifers using a sterile syringe. Both water samples and the rinsed live feed samples were filtered through sterile, hollow fiber syringe filters for aqueous solutions (0.2 μm 2.5 cm^2, DynaGard, Microgon Inc., California) and stored at −20°C. On average 3 cod larvae from each tank were sampled on 8, 17, 30 and 46 dph, by syphoning water through a plastic tube at the middle depth of each tank. Live cod larvae were selected randomly and sacrificed by an overdose of tricaine methanesulfonate (MS222) before further processing. The larvae were rinsed twice in sterilized seawater and transferred individually to Eppendorf tubes, immediately snap frozen in liquid nitrogen and stored at −20°C.

DNA extraction

DNA was extracted using a DNeasy Blood and Tissue Kit (Qiagen). DNA extraction from individual cod larvae, live feed and water samples was performed as described in the protocol for Gram-positive bacteria by the manufacturer, but with minor modifications (for details, see Attramadal et al. 2014).

PCR amplification and sequencing

Fish, feed and water samples were prepared for Illumina MiSeq sequencing by amplification of the V4 region of the 16S rRNA gene using the following primers (locus-specific V4 primer underlined and bold) including 5′ adapter sequences for later indexing PCR and Illumina MiSeq sequencing:

515 F 5′ TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGNNNN **GTGCCAGCMGCCGCGGTAA** 3′

(Caporaso et al. 2011) and

803 R 5′ GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGNNNN**CTACVVGGGTATCTAAKCCBK** 3′. The 803 R primer was designed for this study, because previously published PCR primers targeting this region appear to target and co-amplify algal chloroplast 16S rDNA. Since *Nannochloropsis oculate* algal paste (ReedMariculture) was used in the fish tanks, preliminary PCR and subsequent sequencing for water samples revealed that co-amplification of *Nannochloropsis* 16S rDNA was a major problem. Alignment of *Nannochloropsis oculate* chloroplast and bacterial 16S rRNA gene sequences were used to identify bacteria-specific sequences in the same gene region, and the RDP Probe-match tool was used to examine coverage among bacteria.

To obtain approximately the same amount of PCR product for all samples, the reactions were run for 38 cycles for water samples and 40 cycles for cod larval samples (98°C 15 s, 55°C 20 s, 72°C 20 s) with 0.3 μM of each primer, 0.25 mM of each dNTP, 1 mM $MgCl_2$, 12 μM of bovine serum albumin (BSA), glycerol (10%), Phusion Hot Start II High-Fidelity DNA Polymerase and reaction buffer from Thermo Scientific in a total volume of 20 μL. PCR products were evaluated on a 1% agarose gel, and purified and normalised using a SequalPrep™ Normalization Plate Kit (Invitrogen). A second PCR was performed to attach dual indices to the normalised amplicons by using the Nextera XT Index Kit. The reactions were run for 8 cycles (98°C 15 s, 50°C 20 s, 72°C 20 s) with 0.25 mM of each dNTP, Phusion Hot Start II High-Fidelity DNA Polymerase and reaction buffer (Thermo Scientific) and 2.5 μl of each index primers in a total volume of 25 μL. The indexed PCR products were purified and normalised as described above, pooled, and concentrated by using Amicon® Ultra-0.5 Centrifugal Filter Devices. The resulting amplicon library was sequenced on two MiSeq lanes (Illumina, San Diego, CA) employing 260 bp paired-end reads at the Norwegian Sequencing Center at the University of Oslo, Norway. The resulting Illumina sequencing data are deposited at the European Nucleotide Archive (accession numbers ERS4778574–ERS4778759).

DNA sequence data processing

The Illumina sequencing data were processed using USEARCH utility (version 11) (http://drive5.com/usearch/features.html). The command Fastq_mergepairs was used for merging of paired reads, trimming off primer sequences and filtering out reads shorter than 230 base pairs. The processing further included demultiplexing, removal of singleton reads, and quality trimming (the Fastq_filter command with an expected error threshold of 1). Chimera removal and clustering at the 97% similarity level was performed using the UPARSE-OTU algorithm (Edgar 2013). Microbial taxonomy assignment was performed applying the Sintax script (Edgar 2016) with a confidence value threshold of 0.8 and the RDP reference data set (version 16) (Mollerup et al. 2016). OTUs (operational taxonomic units) of particular interest were further analysed with the SINA tool at the SILVA web site (www.arb-silva.de). OTUs representing algae, Archaea and *Cyanobacteria/Chloroplast* were removed from the OTU table. An OTU found to represent *Propionibacterium acne*, a well-known contaminant of DNA extraction kits (Mollerup et al. 2016) was removed. To remove biases due to variation in sequencing depth, statistical analyses were performed on an OTU table that had been subsampled to 12100 sequencing reads for each sample (the threshold was chosen based on the sample with the lowest number of reads).

Statistical analysis

Ordination by principal coordinate analysis (PCoA) based on Bray–Curtis similarities (Bray and Curtis 1957) was used to visualise differences in microbial community composition between groups of samples. One-way and two-way PERMANOVA (Anderson 2001) based on Bray–Curtis similarities were used to test for statistically significant differences in microbial community composition between groups of samples. Similarity percentage analysis (SIMPER) was used to identify OTUs responsible for

differences (measured as Bray–Curtis similarities) between different groups of samples. Ordination, PERMANOVA (Permutational Analysis of Variance) and SIMPER were performed using the program package PAST version 3.22 (Hammer, Harper and Ryan 2001). Venn diagrams were created using an online tool from Ghent University (http://bioinformatics.psb.ugent.be/webtools/Venn/). Alpha diversity was evaluated as Hill numbers (Tuomisto 2012) with the *reyni* function from the *vegan* package in R (https://cran.r-project.org/web/packages/vegan/index.html, version 2.5–6). Richness, or diversity of order 0, counts the number of OTUs in each sample, while evenness was defined as diversity of order 1 divided by diversity of order 0. MEGA X software (Kumar *et al.* 2018) was used to align OTU sequences by the Muscle algorithm and to make a neighbour-joining phylogenetic tree of the OTUs in the dataset.

Estimation of significance of ecological processes

To evaluate the bacterial community assembly in individual fish and water samples, the nearest taxon index (NTI) was calculated, as described by Webb (2000). NTI values >2 indicate that the community is more phylogenetically clustered than expected by chance, and thus that deterministic processes such as environmental selection have structured the community (Zhou and Ning 2017). NTI was calculated for each water and fish sample with the function ses.mntd(null.model = 'taxa.labels', abundance.weighted = FALSE) in the R-package picante (version 1.8) (Kembel *et al.* 2010). To investigate the ecological processes causing inter-individual variation in the cod larval communities, we used a null-model based statistical framework based on β-NTI, which reflects the phylogenetic dissimilarity between two microbial communities (Stegen *et al.* 2015). β-NTI was calculated within each system and sampling day. With β-NTI values <-2, the communities have similar phylogenetic clustering patterns, and it is assumed that community assembly is caused by homogenous selection. β-NTI values >2 indicate that the communities are under heterogenous selection pressure. For β-NTI values not significantly different from the null model (i.e. $|\beta\text{-NTI}| < 2$), the community assembly is primarily assumed to be driven by stochastic processes, such as drift, homogenising dispersal and dispersal limitation acting with drift.

RESULTS

Atlantic cod larvae were reared with three different microbial metacommunities resulting from three independently operated water treatment systems: a traditional FTS, an MMS and an RAS. After quality trimming and chimera removal, we obtained 2 378 168 sequence reads from 16S rRNA amplicon sequencing of water, feed and fish samples. A total of 3371 OTUs were detected after subsampling to an equal number of amplicon reads (12 100) for each sample.

Alpha diversity

Comparing the estimated number of OTUs (Chao1; Supplementary Table S2, see online supplementary material), with the observed OTU richness (Supplementary Fig. S1, see online supplementary material) revealed that the sequencing depth covered approximately 85, 67 and 75% of the estimated total richness for fish, tank water and feed samples, respectively. The observed OTU richness in tank water typically exceeded that in young larvae (8 dph) by a factor 2.5, while at 46 dph it was only 1.8 times higher. The richness was more stable for tank water bacteria than for fish bacteria in all rearing systems throughout the experiment. Interestingly, for fish bacteria both the observed richness and the evenness (Supplementary Fig. S1) was lowest at the two earliest sampling points and increased with age. This was particularly pronounced for the evenness, which increased more than 10 times from 8 to 46 dph in FTS. In RAS, the evenness was higher at 8 dph than at 17 and 30 dph.

Environmental bacterial communities

The source bacterial community included the bacterial communities in the incoming water, in the tank water, associated with the feed and associated with the fish at the time of transfer to the rearing tanks. All systems received water from the same source and feed from the same cultivation tank. Thus, all OTUs detected can be considered a global species pool. Ecological processes within the three systems further structured the three metacommunities.

The incoming bacteria and selection structured the tank water communities

The different water treatment systems yielded significantly different bacterial communities in the water going into the fish tanks (Bray–Curtis similarity indices, one-way PERMANOVA, $P \leq 0.03$). The bacteria of the incoming water clearly affected the composition of the tank water communities in all systems (Fig. 1). The communities in the incoming water and the tank water was similar in RAS (Fig. 1), whereas for both MMS and FTS the communities of the incoming water differed significantly from those of the tank water (Bray–Curtis similarity indices, one-way PERMANOVA, $P < 0.002$). This indicates that the microbes in the water were under different selection pressures in the FTS and MMS tanks compared to the incoming water in these systems. The tank water communities in RAS differed significantly from those in FTS and MMS (Bray–Curtis similarity indices, one-way PERMANOVA, $P = 0.0001$). However, there were no significant differences in tank water communities between FTS and MMS ($P = 0.7$). At 46 dph, when all systems had received identical incoming water (MMS) for 2 weeks, the bacterial communities of the tank water appeared to be more similar between RAS and FTS/MMS than earlier in the experiment (Fig. 1). This was corroborated by average Bray–Curtis similarities (Supplementary Fig. S2, see online supplementary material), with significantly higher similarity between RAS and FTS tank water, and RAS and MMS tank water at 46 dph than at 30 dph (t-test, $P = 0.002$ and 0.0004, respectively). This indicates that both the composition of the bacterial communities in the incoming water and the selection in the fish tanks had an impact on the composition of the tank water communities. Moreover, the water communities in each replicate tank was more phylogenetically structured than expected by stochastic assembly (NTI > 2), indicating that they were assembled by deterministic processes such as selection (Supplementary Fig. S3, see online supplementary material).

At the OTU level, the bacterial community composition of the tank water varied considerably over time and between systems (Supplementary Fig. S4, see online supplementary material). As an example, OTU_6 (classified as *Leucothrix* by the SILVA database classification tool) dominated in the FTS and MMS tank water at 8 dph, accounting for around 30% of the reads. Its relative abundance decreased dramatically at 30 dph. In RAS, OTU_6 was low in abundance throughout the experiment (~0.25% of the reads). Other examples are a high abundance of an *Aliivibrio* OTU (OTU_28), exclusively in the FTS tank water at 30 dph

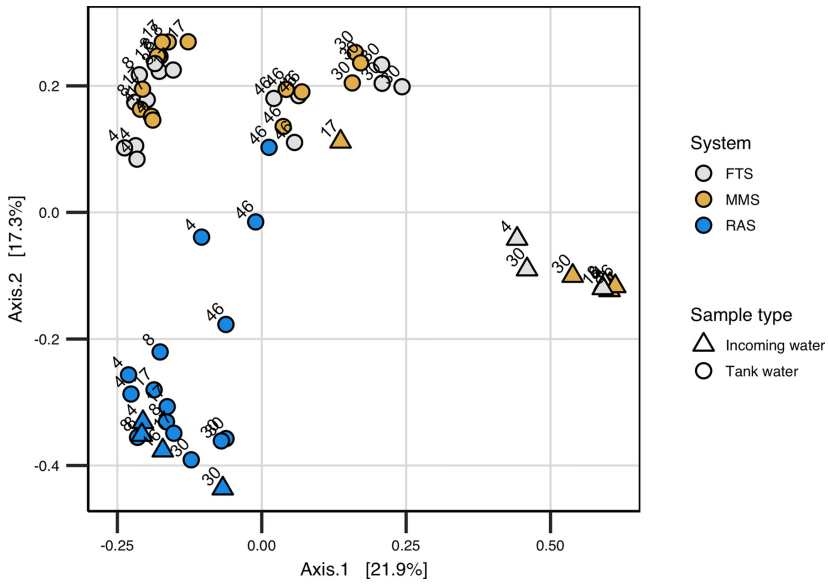


Figure 1. PCoA ordination plot based on Bray–Curtis similarities for comparison of the bacterial communities of tank water and incoming water at 4, 8, 17, 30 and 46 dph (day 4–day 46) in FTS, MMS and RAS. All tanks received MMS water from 31 dph onwards. Colours indicate the water treatment system, numbers in plot indicate the sampling-day and shapes the sample type (triangle = incoming water, circle = tank water).

(average 21% of the reads), and a predominance of a *Polaribacter* OTU (OTU_14) in RAS tank water, with maximum abundance at 30 dph (~40% in two of the replicate tanks).

The bacterial comminities associated with the feed were influenced by the tank water communities

All live feed distributed to the fish tanks originated from the same cultivation tank. Thus, we assumed that the bacterial communities of the feed introduced to the systems were identical. However, rotifers actively ingest bacteria, and the microbial communities may change before the feed is eaten by the fish. The bacterial community of the feed sampled from the RAS rearing tanks differed significantly from those in FTS and MMS rearing tanks (one-way PERMANOVA; $P = 0.015$ and 0.0033 for comparison with FTS and MMS, respectively). There were no significant differences in the bacterial communities of the feed between the FTS and the MMS ($P = 0.8$, Fig. 2A), which indicates that the tank water communities influenced the feed communities. However, the tank water and feed communities differed significantly in each of the systems at each sampling time (one-way PERMANOVA, $P < 0.02$), indicating that either selection or dispersal were dominating the community assembly of the live feed. The Bray–Curtis similarity between the feed and water samples was approximately three times lower in RAS compared with FTS and MMS (Fig. 2A and B).

Selection in the host contributed to the bacterial community assembly of the fish

The bacterial communities of the fish were highly dissimilar from those of the tank water and live feed throughout the experiment, as reflected by both PCoA ordination (Fig. 3A) and the community composition at the order level (Fig. 4). Average Bray–Curtis similarities for water/feed vs fish comparisons within systems and sampling times ranged from 0.0013 to 0.22 (including standard deviations in Fig. 3B and C). The differences were significant for all systems at all sampling times (one-way PERMANOVA, $P < 0.006$ and $p \leq 0.005$ for water and feed comparisons, respectively). Interestingly, as much as 63% of all OTUs observed in the fish were unique for fish samples (Supplementary Fig. S5, see online supplementary material). Moreover, NTI values for individual fish communities indicated more phylogenetic clustering than expected for stochastic community assembly (Supplementary Fig. S3). This implies that selection was important for community assembly within the fish.

Stochastic processes contributed to variation in the bacterial communities between individual cod larvae

Average Bray–Curtis similarities show that the fish bacterial communities varied among individuals in the same rearing tank and between replicate tanks (i.e. within treatments), and especially at 30 and 46 dph (Fig. 5A and B, respectively). This indicates that processes such as drift or heterogenous selection in the individual larvae and rearing tanks also contributed to the community assembly in the fish. The β-NTI analysis indicated that there was a temporal increase in the contribution of stochastic processes to the difference in community structure between individual fish (Fig. 6 and Supplementary Fig. S6, see online supplementary material). This increase suggests that selection within the hosts was most important during the early rearing life stages, and that stochasticity created variation in the communities between fish in the same system and with increasing importance over time. This temporal trend was most pronounced in RAS, with a gradual increase in the relative contribution of stochastic processes from 41 to 75%. None of the comparisons were categorised as heterogeneous selection (β-NTI > 2), indicating that the selection pressure was similar

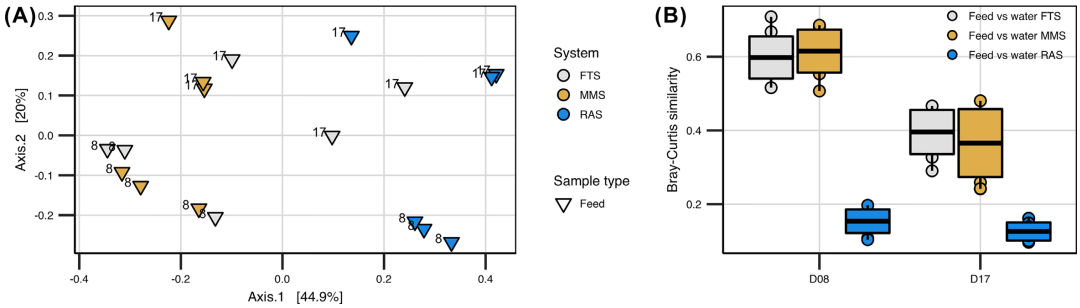


Figure 2. Comparison of the composition of the bacterial communities in feed and water samples from the tanks for the different water treatment systems (FTS, MMS and RAS) taken at 8 and 17 dph. **(A)** PCoA ordination plot based on Bray–Curtis similarities. Colours indicate the water treatment system and numbers in plot the sampling-day. **(B)** Bray–Curtis similarities of the bacterial communities between feed and water within each system at each sampling-time. Each box is based on nine comparisons (three water and three feed samples). Solid black line indicates mean similarity and the surrounding box the standard deviation.

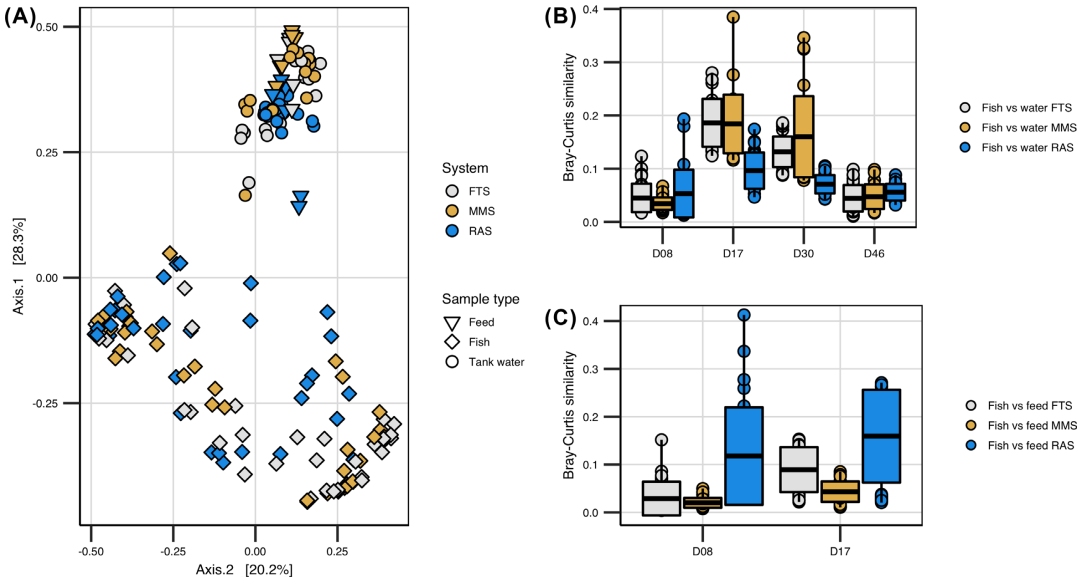


Figure 3. Comparison of bacterial communities of fish, water and live feed samples. **(A)** PCoA ordination plot based on Bray–Curtis similarities for comparison of bacterial communities of all fish, water and feed samples in all treatment systems throughout the experiment. Colours indicate the water treatment system and shapes the sample type (triangle = feed, diamond = fish larvae and circle = tank water). **(B)** Bray–Curtis similarities for comparison of fish and tank water bacterial communities in FTS, MMS and RAS at 8, 17, 30 and 46 dph. Each box is based on 27 comparisons (three water and nine fish samples). Solid black line indicates mean similarity and the surrounding box the standard deviation. **(C)** Bray–Curtis similarities for comparison of communities of fish and feed samples in FTS, MMS and RAS at 8 and 17 dph. Each box is based on 27 comparisons (three feed and nine fish samples). Solid black line indicates mean similarity and the surrounding box the standard deviation.

among individual hosts, and that the inter-individual variation was due to stochastic processes.

The bacterial communities of the water affected the bacterial communities of the fish

Despite the high dissimilarity between bacterial communities in tank water and fish, the fish communities differed according to rearing system. This indicates that the bacteria in the water influence the bacterial communities of the fish. The fish communities in RAS differed significantly from those in FTS and MMS at 8, 17 and 30 dph (one-way PERMANOVA $P < 0.03$) (Fig. 7A–C), with average Bray–Curtis similarities of around 0.4–0.5 for between-system comparisons (Supplementary Fig. S7, see online supplementary material). For FTS and MMS, however, the fish communities were more similar (average Bray–Curtis similarities of 0.6–0.9, Supplementary Fig. S7), and differed significantly only at 30 dph (one-way PERMANOVA, $P = 0.02$). Interestingly, at 46 dph, when all fish tanks had received the same incoming water (MMS) for 16 days, there were no significant differences in the fish communities between any of the systems (Fig. 7D). Moreover, the average Bray–Curtis similarities were comparable within and between systems (Fig. 5 and Supplementary Fig. S7). This clearly points to an influence of the environmental bacteria on the fish communities. Next, we investigated the metacommunities in more detail to elucidate the influence

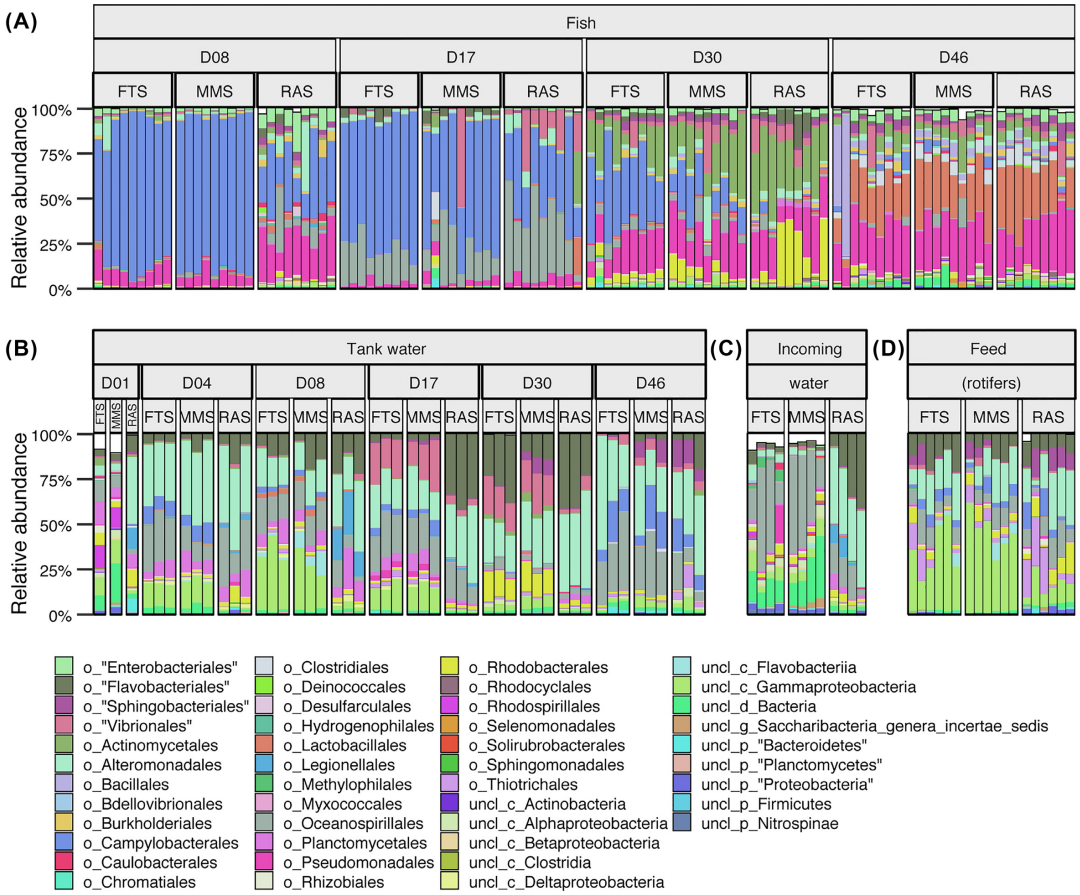


Figure 4. Relative abundances of the 45 most abundant bacterial orders detected in **(A)** fish, **(B)** tank water, **(C)** incoming water and **(D)** feed samples in the systems FTS, MMS and RAS; D01-D46 indicates sampling times given as dph.

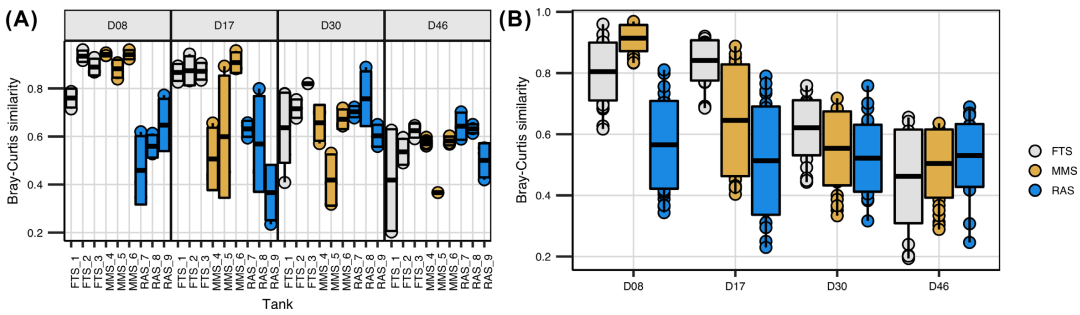


Figure 5. Bray–Curtis similarities for comparisons of bacterial communities of individual fish within **(A)** and between **(B)** replicate rearing tanks (FTS1–3, MMS4–6 and RAS7–9) at 8, 17, 30 and 46 dph (D8–D46). Comparisons are based on between two and four individuals from each tank (A) or nine samples from each system and sampling time (B). For FTS at 30 dph and MMS at 46 dph, only two cod larvae were sampled from one of the tanks, but the total number of fish sampled from each system was always nine. Solid black line indicates mean similarity and the surrounding box the standard deviation.

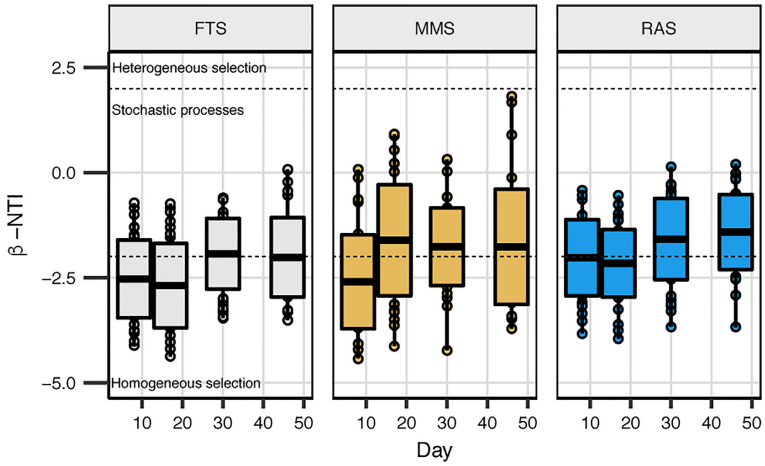


Figure 6. β-NTI for comparisons of fish bacterial communities between individuals within each system (FTS, MMS and RAS) at each sampling time. β-NTI > 2 and < -2 indicates heterogeneous and homogeneous selection, respectively. $|\beta\text{-NTI}| < 2$ represents comparisons that are not significantly different from the null model and indicate stochastic community assembly. Solid black lines indicate the mean β-NTI value on a sampling day within a treatment system ($n = 36$), and the surrounding box the standard deviation.

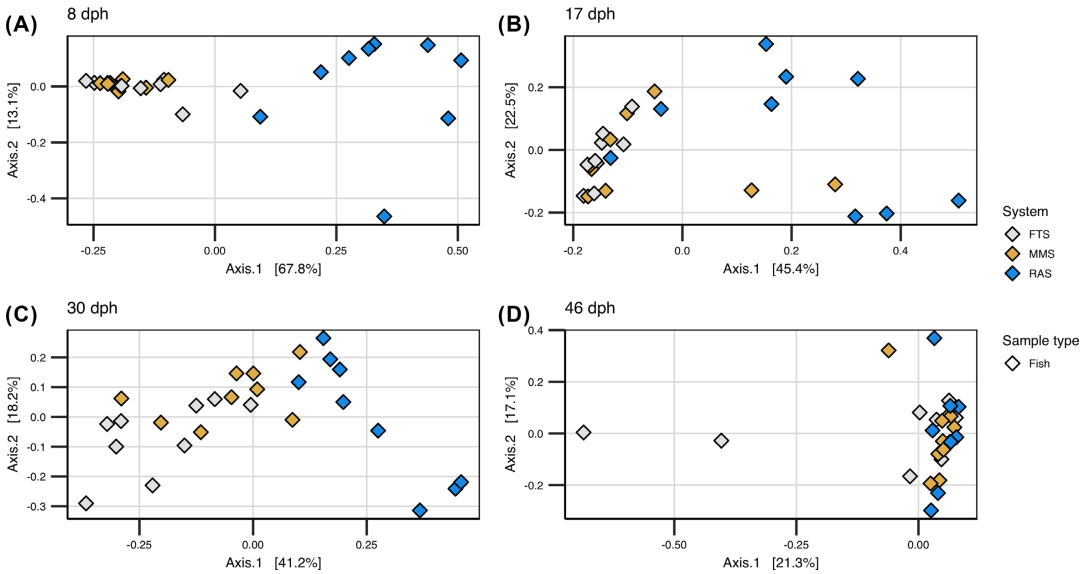


Figure 7. PCoA ordination plot based on Bray–Curtis similarities for comparison of fish bacterial communities in FTS, MMS and RAS at **(A)** 8, **(B)** 17, **(C)** 30 and **(D)** 46 dph. Colours indicate water treatment system.

of the water bacteria on the fish communities at the OTU level. Correlating the number of reads for each OTU in both fish and tank water samples (Supplementary Fig. S8a-c, see online supplementary material), revealed that only five OTUs in the whole data set reached average abundances larger than 2% (of the total reads in at least one sample) for both fish and water samples in at least one system and sampling time (Supplementary Table S3, see online supplementary material). This implies that distinct selection regimes act on the water and the fish bacteria, and that few bacterial populations were selected for in both environments. Only OTU_3 (*Marinomonas*) and OTU_13 (*Aliivibrio*) were more abundant than 5% in both sample types (Supplementary Table S3). The *Marinomonas* OTU was abundant in both water and fish samples at 17 dph in all systems. A SIMPER analysis identified the OTUs contributing most to the differences in fish communities between RAS and FTS/MMS and we identified the average abundance of these OTUs in the relevant water samples (Supplementary Table S4, see online supplementary material). An OTU representing *Arcobacter* (OTU_1) contributed most to the dissimilarity at both 8, 17 and 30 dph (explaining 46, 33 and 21% of the Bray–Curtis dissimilarity, respectively). This OTU dominated the fish samples in MMS and FTS at 8 and 17 dph (average

81 and 65% of the total reads, respectively), and was almost 20 times more abundant in FTS/MMS larvae than in RAS larvae at 30 dph. It was also far more abundant in the water in FTS and MMS compared with RAS at 8 and 17 dph (∼40 times), but the maximum abundance never exceeded 2.3% on average in the water of any system on these days. Thus, even though highly distinct bacterial communities were selected for in water and fish, differences in the relative abundance of rare water OTUs seemed to have a major impact on the bacterial communities of the fish. This supports the above-mentioned conclusion regarding the significance of selection for bacteria associated with the fish.

DISCUSSION

In this study, we used cod larvae as the model organism and investigated the bacterial community assembly during a first feeding experiment. The fish were reared with three different metacommunities (FTS, MMS and RAS) for 30 days, followed by a period of 16 days where all tanks received the same microbiota (MMS). We aimed at quantifying the relative importance of the four high-level ecological processes described by Vellend (2016) and at elucidating the relationship between the bacteria in water and associated with the fish.

Different selection regimes resulted in different bacterial communities in the water

The water treatments differed between the three systems and yielded three incoming waters with distinct bacterial communities. However, the communities in FTS and MMS tank water were not significantly different, but there was variation at the OTU level (Supplementary Fig. S4). In these systems the carrying capacity of the incoming water was higher than in the fish tanks and this promotes fast-growing opportunistic bacteria (Attramadal *et al.* 2014; Vadstein *et al.* 2018). This opportunistic selection most likely caused the similar community composition in the two systems (Vadstein *et al.* 2018). However, at 32 dph the survival of cod in MMS was ∼65% higher than in FTS (results calculated by and presented in Attramadal *et al.*, 2014). Except for the metacommunity composition, the FTS and MMS fish were reared under equal conditions. Therefore, detrimental fish–microbe interactions are the most likely explanation for the difference in survival. In RAS systems, on the other hand, the carrying capacity is similar throughout the system. This restricts opportunistic growth in the fish tanks, and consequently the microbiota in the incoming water and the tank water are similar, as shown by Attramadal *et al.* (2014). Consumption of dissolved organic matter mainly in the biofilters under strong competition, long hydraulic retention time and absence of disinfection created K-selection in RAS water (MacArthur and Wilson 1967; Attramadal *et al.* 2014; Vadstein *et al.* 2018). Thus, both the composition of the incoming water microbiota and the selection forces in the fish tanks contributed to the bacterial community assembly in the tank water.

Bacterial community assembly in cod larvae was dominated by selection and drift

We showed that the bacterial communities of the fish were highly dissimilar from the bacterial communities of the water and feed, indicating that selection was important for community assembly in the fish (Yan, van der Gast and Yu 2012; Bakke *et al.* 2015; Yan *et al.* 2016). This was validated as all NTI values indicated strong phylogenetic clustering and that 63% of the OTUs were unique to the fish. Because all β-NTI values were <2 we ruled out heterogeneous selection as the dominating process steering the bacterial community assembly of the fish. While many analyses indicated strong selection, we observed considerable inter-individual variations both within tanks and between tanks in the same system and the variations seemed to increase with larval age (Figs. 5 and 6). While homogeneous selection should reduce variation, stochastic processes such as dispersal, diversification and drift introduce a randomness that increases variation. Based on traditional metacommunity theory, microbes disperse from the water to the fish and from the fish to the water through excretion (Miller, Svanbäck and Bohannan 2018). Through water, the microbes can disperse from one fish to another given that the fish are in the same tank. In this experiment the major source pools of bacteria to the systems were the incoming water and the feed. The tank water and feed microbiota are the two primary sources of bacteria for fish. Marine fish larvae actively take up bacteria from the surrounding water at rates 100 times higher than the drinking rate, resulting in a consumption of 10^4–10^6 bacteria per larva per day (Reitan, Natvik and Vadstein 1998; Vadstein *et al.* 2018). Ingestion of feed provides an additional 10^5–10^7 (Reitan, Natvik and Vadstein 1998). Consequently, it is unlikely that there was dispersal limitation from the environment to the larvae. Given the length of the experiment we believe it is unlikely that diversification played a role in community assembly at the OTU level (Burns *et al.* 2016). We therefore argue that the processes we have classified as stochastic are equal to drift. β-NTI calculations showed that there was no heterogeneous selection, thus indicating that drift was the main driver of the inter-individual variation observed between larvae. Based on our findings it appears that selection had a major role in structuring the metacommunity, while drift created variation within it. On average, selection within the fish and drift contributed equally to the bacterial community assembly in the cod larvae (Supplementary Fig. S6).

Investigations on community assembly in fish larvae have been done previously (Yan, van der Gast and Yu 2012; Burns *et al.* 2016; Yan *et al.* 2016) and the methods used are either based on composition (neutral model) or phylogeny (β-NTI null-model). In a study on zebrafish, Burns *et al.* (2016) based their estimations on composition, and they argue, as we do, that stochastic processes generate considerable inter-individual variation. However, their results showed that the contribution of stochasticity decreased with host age. In other studies (Yan, van der Gast and Yu 2012; Yan *et al.* 2016) where phylogenetic-based models have been used, as has been done in our study, the results have shown that the contribution of stochasticity increases with host age. Different microbial species in communities may result in large differences between communities when using the neutral model, however if the comparison is based on phylogenetic models the differences might be smaller if the phylogenetic distance between the species is short. This might contribute to explaining the different conclusions drawn by Burns and co-workers and by us. While Burns *et al.* (2016) included the whole life cycle of the zebrafish in their study, we only examined the larval life stage of Atlantic cod. The seemingly contradictory results might also reflect biological significant differences.

Rare OTUs in the bacterial communities of the water may have large consequences for the community assembly of the fish

We have previously shown that it is possible to promote mutualistic fish–microbe interactions in aquaculture systems through well-designed water treatment based on ecological principles (Vadstein *et al.* 2018). In concordance with our previous study (Vestrum *et al.* 2018), we demonstrated here that the bacterial communities of the tank water significantly affected the communities of the fish. From 31 dph onwards, all fish were exposed to the same species pool, and we observed that the differences in the bacterial communties of the fish were reduced between systems. This indicates that the water microbiota affected the bacterial communities of the fish even after the initial colonisation had resulted in different bacterial communities of the fish in the different systems. However, at this point the fish microbiota had probably not yet reached maturity. Therefore, we cannot rule out that other factors such as the developmental stage and the change of feed at 31 dph affected the succession of the bacterial communities as well. The immaturity of the microbiota might also explain the evenness and richness increasing with age. The influence of the water microbes on the bacterial communities of the fish seems to be in contrast with the large dissimilarity observed between the bacterial communities of the fish and the water. The experimental design in this study, including three different microbial water qualities with triplicate tanks, and detailed characterisation of bacterial communities in both fish and water, allowed us to investigate this paradox. Most OTUs found in the fish had low abundances in the water, and vice versa, indicating distinct selection regimes for these two environments. However, a few OTUs were present in relatively high abundances in both water and fish, suggesting that minor fractions of the bacteria were able to compete in both environments. Moreover, we found that an OTU representing *Arcobacter* (a potential opportunistic pathogen (Fitzgerald and Nachamkin 2015) was responsible for most of the differences observed between the bacterial communities of fish in RAS and FTS/MMS. This OTU constituted as much as 81 and 65% of the total reads for FTS and MMS fish samples, respectively, at 8 and 17 dph. The survival of cod larvae was 40% lower in the FTS than in MMS and RAS. This, as well as the excessive dominance of the *Arcobacter* OTU in FTS and MMS larvae, might be explained by differences in the structure of the environmental bacterial communities between the systems, with higher fractions of opportunists in FTS/MMS, and the resulting implications for microbe–microbe and microbe–host interactions. As proposed by Miller *et al.* (2018), the host health might have been affected by the microbiota. In our study, the fish reared in FTS and MMS may have become more susceptible to colonisation by *Arcobacter* due to detrimental host–microbe interactions resulting from the presumably higher fraction of opportunistic bacteria in the rearing water in these systems than in RAS. This shows that the systems' metacommunity should be considered when investigating the community assembly in hosts. Moreover, this study suggests that it is possible to steer the metacommunity towards mutualistic host–microbe interactions.

Sequencing of 16S rDNA amplicons and data processing involving OTU clustering has been the golden standard for microbial diversity studies. However, new approaches have been developed in recent years. For example, as an alternative to the OTU clustering of similar 16S sequences, the concept of amplicon sequencing variants (ASV) has been introduced (Porter and Hajibabaei 2018). ASV-based studies have been suggested to give more realistic and detailed characterisations of microbial communities compared to OTU-based studies (Porter and Hajibabaei 2018). It would be interesting to investigate how this would influence studies on ecological processes in microbial ecology, with the potential for reflecting the actual microbial diversity to a greater extent.

Through a detailed characterisation of the bacterial communities of cod larvae and their environment we were able to elucidate the relationship between host and environmental bacteria. In aquaculture, cod larvae live in a microbial metacommunity that receives bacteria from incoming water and feed. This metacommunity was strongly structured by selective forces, but drift created variation. We were able to identify a single OTU that was selected for in both FTS and MMS and was highly abundant in the fish microbiota in these systems. This OTU might have influenced the survival of the larvae. These findings suggest that it is possible to steer the metacommunity towards mutualistic host–microbe interactions.

SUPPLEMENTARY DATA

Supplementary data are available at *FEMSEC* online.

FUNDING

This study was supported partly by the Faculty of Natural Sciences at the Norwegian University of Science and Technology through scholarships to R.I.V. and M.S.G. and partly by The European Community's Seventh Framework Programme (FP7/2007-2013) [grant number 227197].

Conflict of Interest. None declared.

REFERENCES

Anderson MJ. A new method for non-parametric multivariate analysis of variance. *Austral Ecol* 2001;**26**:32–46.

Attramadal KJK, Truong TMH, Bakke I *et al.* RAS and microbial maturation as tools for K-selection of microbial communities improve survival in cod larvae. *Aquaculture* 2014;**432**: 483–90.

Bakke I, Coward E, Andersen T *et al.* Selection in the host structures the microbiota associated with developing cod larvae (*Gadus morhua*). *Environ Microbiol* 2015;**17**: 3914–24.

Bonder MJ, Kurilshikov A, Tigchelaar EF *et al.* The effect of host genetics on the gut microbiome. *Nat Genet* 2016;**48**:1407.

Boutin S, Sauvage C, Bernatchez L *et al.* Inter individual variations of the fish skin microbiota: host genetics basis of mutualism? *PLoS One* 2014;**9**:e102649.

Bray JR, Curtis JT. An ordination of the upland forest communities of southern wisconsin. *Ecological Monographies* 1957;**27**:325–49.

Brugman S, Ikeda-Ohtsubo W, Braber S *et al.* A comparative review on microbiota manipulation: lessons from fish, plants, livestock, and human research. *Frontiers in Nutrition* 2018;**5**:80.

Burns AR, Miller E, Agarwal M *et al.* Interhost dispersal alters microbiome assembly and can overwhelm host innate immunity in an experimental zebrafish model. *Proc Natl Acad Sci* 2017;**114**:11181.

Burns AR, Stephens WZ, Stagaman K *et al.* Contribution of neutral processes to the assembly of gut microbial communities in the zebrafish over host development. *The Isme Journal* 2016;**10**:655.

Caporaso JG, Lauber CL, Walters WA *et al.* Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc Natl Acad Sci* 2011;**108**: 4516–22.

David LA, Maurice CF, Carmody RN *et al.* Diet rapidly and reproducibly alters the human gut microbiome. *Nature* 2013;**505**:559.

Edgar RC. SINTAX: a simple non-Bayesian taxonomy classifier for 16S and ITS sequences. *bioRxiv* 2016:074161.

Edgar RC. UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nat Methods* 2013;**10**:996.

Fitzgerald C, Nachamkin I. *Campylobacter* and *Arcobacter*. *Manual of Clinical Microbiology, Eleventh Edition*. American Society of Microbiology, 2015, 998–1012.

Fjellheim AJ, Playfoot KJ, Skjermo J *et al.* Inter-individual variation in the dominant intestinal microbiota of reared Atlantic cod (*Gadus morhua* L.) larvae. *Aquaculture Research* 2012;**43**:1499–508.

Fujimura KE, Demoor T, Rauch M *et al.* House dust exposure mediates gut microbiome *Lactobacillus* enrichment and airway immune defense against allergens and virus infection. *Proc Natl Acad Sci* 2014;**111**:805.

Giatsis C, Sipkema D, Smidt H *et al.* The impact of rearing environment on the development of gut microbiota in tilapia larvae. *Sci Rep* 2015;**5**:18206.

Hammer Ø, Harper D, Ryan P. PAST-palaeontological statistics, ver. 1.89. *Palaeontologia Electronica* 2001;**4**.

Hanson CA, Fuhrman JA, Horner-Devine MC *et al.* Beyond biogeographic patterns: processes shaping the microbial landscape. *Nat Rev Microbiol* 2012;**10**:497–506.

Hiippala K, Jouhten H, Ronkainen A *et al.* The Potential of gut commensals in reinforcing intestinal barrier function and alleviating inflammation. *Nutrients* 2018;**10**.

Jackson MA, Bonder MJ, Kuncheva Z *et al.* Detection of stable community structures within gut microbiota co-occurrence networks from different human populations. *PeerJ* 2018;**6**:e4303.

Kembel SW, Cowan PD, Helmus MR *et al.* Picante: R tools for integrating phylogenies and ecology. *Bioinformatics* 2010;**26**: 1463–4.

Koppel N, Maini Rekdal V, Balskus EP. Chemical transformation of xenobiotics by the human gut microbiota. *Science* 2017;**356**:eaag2770.

Kumar S, Stecher G, Li M *et al.* MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol Biol Evol* 2018;**35**:1547–9.

Lazado CC, Caipang CMA, Brinchmann MF *et al.* In vitro adherence of two candidate probiotics from Atlantic cod and their interference with the adhesion of two pathogenic bacteria. *Vet Microbiol* 2011;**148**:252–9.

Lazado CC, Caipang CMA. Probiotics–pathogen interactions elicit differential regulation of cutaneous immune responses in epidermal cells of Atlantic cod *Gadus morhua*. *Fish Shellfish Immunol* 2014;**36**:113–9.

Leibold MA, Holyoak M, Mouquet N *et al.* The metacommunity concept: a framework for multi-scale community ecology. *Ecology Letters* 2004;**7**:601–13.

MacArthur RH, Wilson EO. *The Theory of Island Biogeography*. USA: Princeton University Press, 1967.

Miller ET, Svanbäck R, Bohannan BJM. Microbiomes as Metacommunities: Understanding Host-Associated microbes through metacommunity ecology. *Trends in Ecology and Evolution* 2018;**33**:926–35.

Mollerup S, Friis-Nielsen J, Vinner L *et al.* *Propionibacterium acnes*: Disease-Causing agent or common contaminant? Detection in diverse patient samples by Next-generation sequencing. *J Clin Microbiol* 2016;**54**:980.

Naito T, Mulet C, De Castro C *et al.* Lipopolysaccharide from Crypt-Specific core microbiota modulates the colonic epithelial Proliferation-to-Differentiation balance. *mBio* 2017;**8**:e01680–17.

Nayak SK. Role of gastrointestinal microbiota in fish. *Aquaculture Research* 2010;**41**:1553–73.

Nemergut DR, Schmidt SK, Fukami T *et al.* Patterns and processes of microbial community assembly. *Microbiol Mol Biol Rev* 2013;**77**:342.

Porter TM, Hajibabaei M. Scaling up: A guide to high-throughput genomic approaches for biodiversity analysis. *Mol Ecol* 2018;**27**:313–38.

Rawls JF, Mahowald MA, Ley RE *et al.* Reciprocal gut microbiota transplants from zebrafish and mice to Germ-free recipients reveal host habitat selection. *Cell* 2006;**127**:423–33.

Reitan KI, Natvik CM, Vadstein O. Drinking rate, uptake of bacteria and microalgae in turbot larvae. *J Fish Biol* 1998;**53**: 1145–54.

Rogers GB, Keating DJ, Young RL *et al.* From gut dysbiosis to altered brain function and mental illness: mechanisms and pathways. *Mol Psychiatry* 2016;**21**:738–48.

Rothschild D, Weissbrod O, Barkan E *et al.* Environment dominates over host genetics in shaping human gut microbiota. *Nature* 2018;**555**:210.

Semova I, Carten Juliana D, Stombaugh J *et al.* Microbiota regulate intestinal absorption and metabolism of fatty acids in the Zebrafish. *Cell Host & Microbe* 2012;**12**:277–88.

Sonnenburg JL, Bäckhed F. Diet–microbiota interactions as moderators of human metabolism. *Nature* 2016;**535**:56.

Stegen JC, Lin X, Fredrickson JK *et al.* Estimating and mapping ecological processes influencing microbial community assembly. *Frontiers in Microbiology* 2015;**6**:370.

Sullam KE, Essinger SD, Lozupone CA *et al.* Environmental and ecological factors that shape the gut bacterial communities of fish: a meta-analysis. *Mol Ecol* 2012;**21**:3363–78.

Tuomisto H. An updated consumer's guide to evenness and related indices. *Oikos* 2012;**121**:1203–18.

Vadstein O, Attramadal KJK, Bakke I *et al.* K-Selection as microbial community management Strategy: A method for improved viability of larvae in aquaculture. *Frontiers in Microbiology* 2018;**9**:2730.

Vadstein O, Attramadal KJK, Bakke I *et al.* Managing the microbial community of marine fish larvae: a holistic perspective for larviculture. *Frontiers in Microbiology* 2018;**9**.

Vellend M. *The theory of ecological communities (MPB-57)*. Princeton University Press, 2016.

Vestrum RI, Attramadal KJK, Winge P *et al.* Rearing water treatment induces microbial selection influencing the microbiota and pathogen associated transcripts of cod (*Gadus morhua*) larvae. *Frontiers in microbiology* 2018;**9**.

Vestrum RI, Luef B, Forberg T *et al.* Investigating fish Larvae-Microbe interactions in the 21st Century: Old questions studied with new tools. In: Yúfera M (ed.). *Emerging Issues in Fish Larvae Research*. Cham: Springer International Publishing, 2018, 1–35.

Webb CO. Exploring the phylogenetic structure of ecological Communities: An example for rain forest trees. *Am Nat* 2000;**156**:145–55.

Yan Q, Li J, Yu Y *et al.* Environmental filtering decreases with fish development for the assembly of gut microbiota. *Environ Microbiol* 2016;**18**:4739–54.

Yan Q, van der Gast CJ, Yu Y. Bacterial community assembly and turnover within the intestines of developing Zebrafish. *PLoS One* 2012;**7**:e30603.

Zhou J, Ning D. Stochastic community Assembly: Does it matter in microbial ecology? *Microbiol Mol Biol Rev* 2017;**81**: e00002–17.

Paper III

Aquaculture rearing systems induce no legacy effects in Atlantic cod larvae or their rearing water bacterial communities

scientific reports

Check for updates

OPEN

Aquaculture rearing systems induce no legacy effects in Atlantic cod larvae or their rearing water bacterial communities

Madeleine S. Gundersen[1]✉, Olav Vadstein[1], Peter De Schryver[2] & Kari Johanne Kihle Attramadal[1]

The microbial rearing quality influences the survival of marine larvae. Microbially matured water treatment systems (MMS) provide a more favourable rearing water microbiome than flow-through systems (FTS). It has previously been hypothesised, but not investigated, that initial rearing in MMS leaves a protective legacy effect in Atlantic cod larvae (*Gadus morhua*). We tested this hypothesis through a crossover 2 × 2 factorial experiment varying the rearing water treatment system (MMS vs FTS) and the microbial carrying capacity (+/− added organic matter). At 9 days post-hatching, we switched the rearing water treatment system. By comparing switched and unswitched rearing tanks, we evaluated if legacy effects had been established in the larvae or their surrounding rearing water bacterial community. We analysed the bacterial communities with flow cytometry and 16S rRNA gene sequencing. We found no evidence that the initial rearing condition left a legacy effect in the communities by evaluating the bacterial community diversity and structure. Instead, the present rearing condition was the most important driver for differences in the rearing water microbiota. Furthermore, we found that MMS with high microbial carrying capacity appeared to seed a stable bacterial community to the rearing tanks. This finding highlights the importance of keeping a similar carrying capacity between the inlet and rearing water. Moreover, we reject the hypothesis that the initial rearing condition leaves a protective legacy effect in larvae, as the larval survival and robustness were linked to the present rearing condition. In conclusion, our results highlight the importance of maintaining a beneficial microbial rearing environment from hatching and throughout the larval rearing period.

Early-stage marine larvae have high mortality and are vulnerable to poor microbial rearing conditions, potentially resulting in infections and gut-dysbiosis[1]. However, beneficial fish-microbe interactions can increase survivability, growth, and resistance to detrimental bacterial colonization[2]. In land-based aquaculture, the fish and its microbiota are influenced by the rearing system conditions[3,4], which can be controlled and managed to optimise fish growth and health[5,6]. Fish are in close contact with their surrounding water[7], and it is now well established that the fish microbiota is influenced by, and changes with, its surrounding water microbiota[3]. The fish microbiome is shaped by many variables, including internal factors such as species, genetics and developmental stage[1], and external factors such as feed, rearing system operation and environmental carrying capacity[8]. For this reason, efforts to manage the fish microbiota, and thereby minimize the impact of harmful microbial interactions, are important to increase the production in marine aquaculture.

The rearing water treatment systems can be operated to select for beneficial host microbes[3,9,10]. Disinfection of the intake water is an essential first line of defence against pathogenetic diseases[3]. However, disinfection reduces the bacterial biomass well below the carrying capacity of the system. This reduction results in an environment favouring the growth of opportunistic, often pathogenic, bacteria that thrive when resources are in surplus[3]. Conventional flow-through aquaculture systems (FTS) typically create environments favouring opportunists[9,11–13]. In FTS, the microbial carrying capacity of the rearing water is considerably higher than in the intake water[9]. This elevated microbial carrying capacity in the rearing tanks is due to an increased organic load from fish feed

[1]Department of Biotechnology and Food Science, NTNU-Norwegian University of Science and Technology, Trondheim, Norway. [2]Laboratory of Aquaculture and Artemia Reference Center, Ghent University, Ghent, Belgium. ✉email: madeleine.gundersen@ntnu.no

and faeces and a high hydraulic retention time (HRT) in tanks during larval rearing. Due to the low bacterial load after disinfection and the high microbial carrying capacity, rapid bacterial regrowth is observed in these environments, which are characterized as unstable, with low bacterial community diversity, a high fraction of opportunists and low biological control[3].

However, by applying ecological theory to manage the microbiota of the rearing tanks, it is possible to select against the opportunistic bacteria[7]. Skjermo et al. 1997 proposed to mature the intake microbial community in a maturing biofilter unit to avoid the rapid regrowth in the rearing tanks[14]. In a maturing biofilter, the bacterial regrowth to the microbial carrying capacity of the intake water occurs under strong competition before entering the rearing tanks[14]. The maturing biofilter is inhabited by bacteria that compete for the incoming resources and therefore develops into a stable community dominated by competition specialists with a reduced risk of opportunistic proliferation. Stable competitive environments are characterised by higher diversity and the potential for higher biological control[3]. Using microbially matured systems (MMS) compared to FTS systems has resulted in increased larval viability of Atlantic halibut (*Hippoglossus hippoglossus*)[14], turbot (*Scophthalmus maximus*)[14] and Atlantic cod (*Gadus morhua*)[9].

Microbial communities are assembled through deterministic and stochastic processes[15], as we previously have shown for the microbiota of Atlantic cod rearing water[16]. Processes that happened in the past can leave deterministic legacy effects in the microbial community[17] and the fish[18,19]. Attramadal et al. (2014) observed higher microbial community stability and Atlantic cod larval survivability in MMS compared to FTS systems[9]. The authors proposed that the increased survivability in MMS was due to a beneficial microbiome initially colonizing the larvae or the rearing water during the first days of rearing[9]. It was further claimed that this legacy effect should persist during larval rearing. However, that experiment was not designed to investigate legacy effects, and thus it has not been tested whether the initial rearing conditions leave a legacy effect.

This study investigated whether the initial rearing condition established legacy effects in Atlantic cod larvae or their rearing water microbiota. We used a 2×2 factorial crossover design with rearing water treatment systems (FTS vs MMS) and microbial carrying capacity (added extra organic matter or not) as the experimental variables. After nine days post-hatching, we switched the inlet water treatment system in half of the rearing tanks. We investigated whether the initial water treatment system left legacy effects in two of the system's biological components: the rearing water bacterial communities and the larvae. By comparing the bacterial communities in the rearing water in switched and unswitched tanks, we investigate if a legacy effect was established in terms of the diversity within each rearing tank (α-diversity) and the structure and taxonomic composition of the communities (β-diversity). We hypothesised that the MMS systems would have higher microbial stability and lower fractions of opportunistic and possibly detrimental species than the FTS systems. We increased the microbial carrying capacity in half of the tanks to evaluate the combined effect of treatment and increased population size of bacteria on the larvae. Similarly, we assessed if legacy effects were established in the larvae by determining if there were differences in the larval weight, robustness, and survival between switched and unswitched rearing treatments. Based on a previous study[9], we hypothesised that the initial larval colonisation in MMS would leave a protective legacy effect in the larvae, resulting in increased survival and stress tolerance compared to the larvae reared in the FTS.

Materials and methods

Experimental design and setup. The experiment had a 2×2 factorial design with the rearing water system and microbial carrying capacity as the two factors and was operated for 20 days post-hatching (DPH). Halfway through the experiment (9 DPH), the inlet water treatment system was switched for half of the rearing tanks by changing the inlet water pipes. Intake water (70 m depth, Trondheimsfjord) was sand-filtered (50 µm) and UV-treated. Half of the 16 rearing tanks (100 L, black, coned bottom) received this water directly and were operated as FTS. For the remaining eight tanks, the intake water was microbially matured in a biofilter (MMS) before entering the rearing tanks[12]. The microbial carrying capacity was manipulated by adding 20 mg/L of organic matter daily directly to each FTS rearing tank (FTS+) and the biofilter serving the MMS rearing tanks (MMS+). The organic matter was a mix of tryptone, peptone and yeast extract (6.67 mg/L each). The tanks with added organic matter were characterized as having a high microbial carrying capacity (+), whereas the others had a low capacity (−). We refer to the rearing tanks that switched water treatment during the experiment as, for example, 'MMS+ to FTS+' to indicate that the tanks received MMS+ water for the first nine DPH before switching to the FTS+ treatment for the rest of the experiment. The carrying capacity was not changed for any of the tanks throughout the experiment.

Rearing regime and biofilter pre-cultivation. The Atlantic cod were reared for 20 DPH. Atlantic cod eggs (Havlandet Marine Yngel AS) were disinfected with glutaraldehyde for 10 min (400 ppm) and rinsed in disinfected seawater for 30 s[20]. The larvae hatched at 90–95-day degrees (°d). The experiment was conducted within the Norwegian animal welfare act guidelines[21]. The Norwegian Animal Research Authority (NARA) approved the facility and this experiment under id 6729. This study is reported according to the ARRIVE guidelines (https://arriveguidelines.org/).

Each experimental tank was stocked with larvae (100 larvae/L) and maintained in darkness until 3 DPH, after which they were kept in continuous light. The tank water exchange rate started at 2 and increased to 4 tank volumes day^{-1} at 8 DPH. A feeding robot (Storvik, Norway) added suspended clay (Vingerling K148, WBB Fucs GmbH, Germany) to the fish tanks (0.1 g L^{-1} day^{-1}) from 1 DPH[22]. Larvae were fed rotifers from 3 DPH and a mix of rotifers and artemia from 18 DPH (Supplementary Table 1).

The two biofilters (267 L) were filled 25% with used Kaldnes carriers K1 (Anox Kaldnes) from the same source biofilter. The biofilters were pre-cultivated to ensure that the biofilm had formed sufficiently and that the

microbial communities had stabilised. Six weeks before hatching the two biofilters were operated as batch at 20 °C and fed every second day with 20 mg/L of the organic matter mix. Four weeks before hatching, water and carriers from the two biofilters were mixed to ensure similar biofilm composition. At the same time, the temperature was lowered to 13 °C, each system refilled with 50 L fresh water, and the flow rate increased to 10 L/h. Onwards, the fed-biofilter (MMS+) was added 20 mg/L of organic matter daily, while the MMS− only received incoming fjord water. Three weeks before hatching, the flow rate was increased to 20 L/h.

Larval growth. The larval growth was quantified by weighing the freeze-dried larvae individually (9–10 larvae per tank at 4, 8 and 12 DPH) or as a pool (3–5 larvae and 5 samples per tank at 2, 12 and 17 DPH). Due to high mortality in the FTS+ tanks, data is lacking from rearing tanks connected to that system at 12, 17 and 18 DPH. Larvae were sacrificed with an overdose of MS222 and rinsed with dH_2O.

Larval stress tolerance. The robustness of the larvae was evaluated as percent survival after exposure to different stress tests on 8, 11 and 17 DPH in two side experiments. The two experiments tested the general stress level of the larvae through a "transfer challenge" and the larvae's resistance to invasion stress through a rearing water "invasion challenge". The transfer challenge can be interpreted as a negative control to the invasion challenge as it only reflects the stress of being transferred from the main rearing tank.

Larvae were harvested by siphoning with silicone hose throughout the tank from one or both rearing tanks in each rearing treatment. An exception was tanks with FTS+ as the initial rearing treatment due to high mortality (see Supplementary Table 2 for subsampling overview). The transfer challenge reflecting the general stress of the larval was conducted on 11 and 17 DPH by simply transferring 10–12 larvae and 100 mL of rearing water from each tank to sterile Nunc culture flasks). The invasion challenge was performed on 8, 11 and 17 DPH. First, we transferred 3.5 L of rearing water to a glass bowl and invaded it with $1.5–2.8 \times 10^6$ *Pseudoalteromonas* CFUs/mL and 2.8×10^4 *Polaribacter* CFUs/mL in a glass bowl (see details below). Next, 2×100 mL of this invaded rearing water was transferred to two sterile Nunc culture flasks and 10–12 larvae were added to the flasks. Thus, we had $n = 1$ and $n = 2$ flasks per subsampled tank for the transfer and invasion challenge, respectively. After 24 h, the survival of the larvae in the flasks was determined.

The two bacteria used as the invaders had previously been isolated from the system on marine agar plates and preserved in 20% glycerol at −80 °C. The bacterial DNA was extracted using ZymoBIOMICS MagBead DNA/RNA extraction kit before the 16S rRNA gene was amplified using the broad coverage PCR primers Eub8F (5′-AGAGTTTGATCMTGGCTCAG-3′) and 1492R (5′-GGTTACCTTGTTACGACTT-3′). The reactions were run for 38 cycles (98 °C 15 s, 55 °C 20 s, 72 °C 20 s) with 0.3 µM of each primer, 0.25 mM of each dNTP, 1× Phusion buffer HF, 0.015 units/µL of Phusion Hot StartII DNA polymerase, 1 µL of DNA template and dH20 to a total volume of 25 µL. The DNA sequences were obtained through Sanger sequencing (LightRun, Eurofins). The bacteria were identified to belong to the *Pseudoalteromonas* and the *Polaribacter* genera, respectively, through the SeqMatch function of the RDP database[23].

Survival in rearing tanks. Survival was quantified at 20 DPH as the percentage of remaining larvae compared to initial stocking in each tank. The remaining larvae at 20 DPH were sacrificed as described above before counting.

Bacterial density and net growth potential. The bacterial density in incoming- and rearing water was quantified using flow cytometry (BD accuri C6) in samples collected at 2, 9, 11 and 15 DPH. Each sample was split into two aliquots; one was fixated immediately with 1% glutaraldehyde and used to quantify the bacterial density. The other was incubated as is in the fish rearing room in cell culture tubes without shaking for three days before fixation. The incubated samples were used to determine the net growth potential of the bacterial community. We calculated the growth potential as the logarithmic (base 2) ratio between the bacterial density in incubated and non-incubated samples. Thus, the net growth potential represents the number of doublings in density after incubation. We defined samples as being at the microbial carrying capacity if the net growth potential was < 0.

16S rRNA gene amplicon library preparation and sequencing. The bacterial communities of the rearing water were filtered through Dynaguard syringe filters (0.2 µm, 50 mL) at 1 and 12 DPH and stored at −80 °C until DNA extraction. DNA extraction and amplicon library preparation was carried out as described in Gundersen et al. 2021[24]. Briefly, bacterial community DNA was extracted using the Qiagen DNeasy PowerSoil DNA extraction kit. Then, broad-coverage primers were used to amplify the V3-V4 region of the 16S rRNA gene using PCR. The amplicon library was then normalised and purified before amplicon indexing with the Illumina Nextera XT Index kits (FC-131-2004). Finally, the amplicon library was sequenced with Illumina MiSeq at the Norwegian Sequencing centre[25]. The sequencing reads are deposited at the European Nucleotide Archive (accession number ERR9837055-ERR9837086). The 16S rRNA gene amplicon dataset contained 450,369 sequence reads with a mean sequencing depth of 14,074 (± 6418 SD) reads per sample.

Processing of Illumina sequence data. The USEARCH pipeline was used to process the Illumina sequence data[26]. First, paired ends were merged simultaneously as primer sequences and reads shorter than 400 bp were removed. Then Unoise3 was used to perform error correction of the amplicon reads, and an amplicon sequence variant (ASV) table was generated[27]. Finally, ASV sequences were taxonomically assigned to the

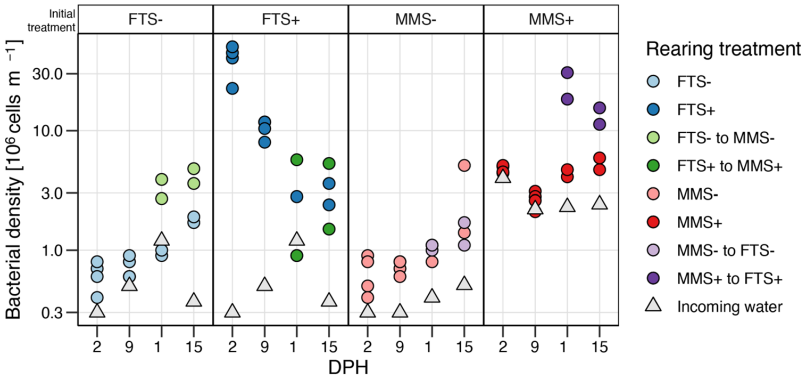


Figure 1. Bacterial density (million bacterial cells mL^{-1}) at various days post-hatching (DPH) in incoming and rearing tank water. Note that the y-axis is log scaled. Colours indicate the rearing treatment, and shape signifies rearing (filled circle) and incoming water (filled triangle).

rdp dataset (rdp16s_v18)[23] version 18 at an 80% confidence level using the sintax command[28] yielding 1315 ASVs.

All data analysis was subsequently performed in R version 4.1.0[29]. First, ASV sequences were multi-aligned using the *AlignSeqs*() function from the DECIPHER package[30]. Then the phangorn package was used to construct a phylogenetic tree from the alignment using neighbour-joining, which was fitted to a generalised time-reversible maximum likelihood tree[31]. All ASVs with less than 8 reads and those identified as non-bacterial were removed from the dataset. Next, the tree was rooted to the longest branch using *root*() from the package ape[32]. Next, each sample was scaled to the lowest sequence depth using *transform*() from the package microbiome[33]. This scaled dataset was rarefied using *rarefy_even_depth*() from phyloseq to ensure equal sampling depth[34]. An assessment of the 16S rRNA gene amplicon dataset quality can be found in Supplementary materials online. All plots were generated using the packages ggplot2[35] and ggpubr[36].

Statistical analysis. The α-diversity was estimated as Hill diversity of order 0 (i.e. richness) and 1 (i.e. exponential Shannon)[37]. These diversity numbers were estimated using the function *reyni*() from vegan[38]. One-way analysis of variance (ANOVA) was used to test for differences in α and β-diversity and larval weight means between groups[39]. The data were tested for homoscedastic variance using the Flinger-Killeen test[40] and for normal distribution with the Shapiro–Wilk's test[41] using the functions *fligner.test*() and *shapiro.test()*, respectively. When the requirements for ANOVA were not met, the Kruskal–Wallis test was used[42]. The Tukey test[43] was used for post hoc comparisons of group means using the function *TukeyHSD()*.

To investigate differences in community composition between samples, we calculated the Bray–Curtis and UniFrac distance and their incidence-based equivalents Sørensen and unweighted UniFrac distance. The distance matrixes were calculated with *distance*(), ordinated with a principal coordinate analysis (PCoA) using *ordinate*() and plotted using *plot_ordination*() from phyloseq[34].

We used DeSeq2 to perform a differential abundance test. DeSeq2 quantifies which ASVs that have significantly different abundances between groups[44]. Briefly, the un-normalized ASV table was used for the DeSeq2 analysis. First, the count data were median ratio normalised using *etimateSizeFactors*(). Then, the dispersion for each ASV was estimated using *estimateDispersion*(). A Wald significance test was then performed on a parametric fitted negative binomial GLM model using *DESeq*(test = " Wald", fitType = "parametric").

Results

Bacterial density and growth potential in the rearing water were related to the microbial carrying capacity. Quantifying the bacterial density in each tank verified that we obtained a higher bacterial load in the systems with added organic material. The bacterial density was, on average, 7.8× higher in the systems with high compared to low bacterial carrying capacity. This difference was particularly evident at 2 (34.8×, Kruskal–Wallis p = 0.0008) and 9 DPH (9.1×, Kruskal–Wallis p = 0.0007) (Fig. 1). The bacterial density increased throughout the experiment for the tanks with low microbial carrying capacity (treatment group MMS−, FTS−), reflecting increased larval feeding and defecation. Contrastingly, the bacterial density was relatively stable over time in the MMS+ treatment and even decreased over time in the FTS+ treatment. When averaging the densities at 11 and 15 DPH within each rearing treatment, we observed that the 'MMS+ to FTS+' had a considerable difference in the bacterial density between incoming and rearing water (24.2×). In contrast, this difference was below 8.2× in all other treatment tanks. Such differences in density indicated that some communities were below the microbial carrying capacity of the systems. We thus investigated the growth potential to determine if carrying capacity was reached in the rearing water.

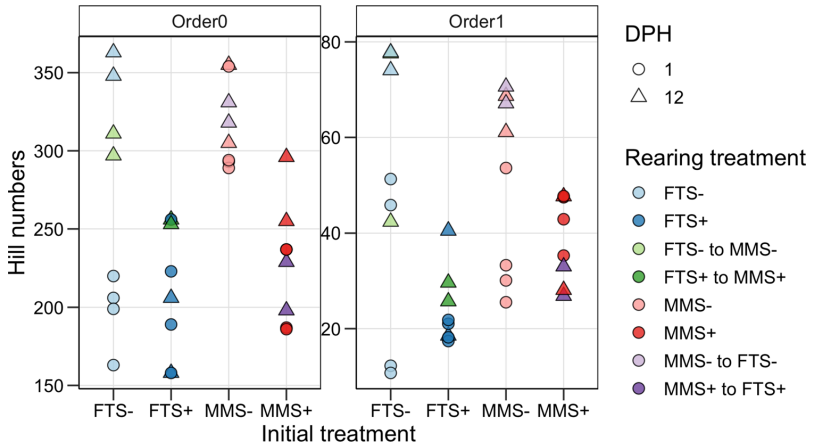


Figure 2. The bacterial α-diversity of Hill diversity orders 0 and 1 at 1- and 12-days post-hatching (DPH). Colours indicate rearing treatment, and shape signifies 1 (filled circle) and 12 DPH (filled triangle). Hill diversity of order 0 is equivalent to ASV richness, and order 1 is equal to exponential Shannon, which also accounts for ASV abundances.

The bacterial net growth potential in the intake and rearing water was quantified as the number of cell doublings after incubation for 3 days[11]. Generally, the FTS− and MMS− rearing water had net growth potential with an average of 0.2 and 0.1, respectively (Supplementary Fig. 2). In contrast, the rearing water of the FTS+ and MMS+ had a negative net growth potential with averages of −0.2 and −0.06, respectively. In the case of negative net growth potential, the bacterial density decreased during the incubation. A negative net growth potential suggested that the rearing water bacterial communities were at the tank's microbial carrying capacity at the time of sampling. Thus, the bacterial communities were at the carrying capacity of the high (+) carrying capacity systems and below in the low (−) systems. To gain a deeper understanding of the bacterial community characteristics the 16S rRNA gene of the bacterial community was sequenced at 1 and 9 DPH.

Initial rearing condition did not leave a legacy effect on bacterial α-diversity. The bacterial α-diversity of the rearing water was investigated at 1 and 12 DPH (Fig. 2). At 1 DPH, the richness was comparable between the FTS−, FTS+ and MMS+ treatments, but on average, 1.5× higher for the MMS− treatment (307 vs 205 ASVs, Tukey's test $p < 0.006$). The diversity of order 1 was, on average, 1.5× higher for the MMS+ and MMS− treatments than for the FTS+ and FTS− treatments (ANOVA $p = 0.05$).

We were interested in determining whether the initial rearing system had a legacy effect on α-diversity. We first evaluated whether there were differences between the unswitched treatments at 12 DPH. For the high carrying capacity treatments, the MMS+ had, on average more ASVs than the FTS+ treatment (275 vs 182 ASVs, Tukey's test $p = 0.04$). For the low carrying capacity group, the MMS− group had, on average fewer ASVs than the FTS- treatment (330 vs 356 ASVs, Tukey's test $p = 0.9$). Note that statistical tests with data from 12 DPH have low power ($n = 2$ replicates/group). Comparing the switched tanks to those that continued with the initial treatment showed that 'FTS− to MMS−', 'FTS+ to MMS+' and 'MMS+ to FTS+' had a more similar richness to their post-switch treatments. Only the 'MMS− to FTS−' treatment had a more similar richness to the initial treatment. However, only 25 ASVs, on average, differentiated MMS− and FTS−. We thus conclude that the initial rearing treatment did not leave a legacy effect on richness. Similarly, there was no indication that the initial rearing treatment had a legacy effect on the diversity of order 1.

However, we did observe that the richness had increased in all treatments, except in the tanks continuing with FTS+. The increase in richness was similar in the tanks with low carrying capacity (FTS− and MMS−) regardless of whether the tank changed water treatment system or not. However, for the tanks with FTS+ as the initial treatment, the richness decreased 0.88× in the tanks continuing with FTS+ but increased 1.2× for tanks that switched to the MMS+ system. Interestingly, the opposite was observed for the tanks starting with MMS+. For these, the richness increased 1.3× in the tanks continuing with MMS+ but was stable for those that switched to FTS+ (1.0×). There were few differences in diversity of order 1 between the switched and unswitched treatments at 12 DPH. However, the diversity of order 1 had increased in all treatments, except in the tanks starting with the MMS+ treatment.

We interpret the increases in α-diversity as indicating that the bacterial communities were unstable at 1 DPH, thus allowing the inlet bacteria to disperse and establish. Notably, the decrease in diversity of order 1 in the tanks starting with MMS+ suggests that these bacterial communities were stable, more even, and resisted the establishment of the microbiota from the new intake water source (e.g. 'MMS+ to FTS+' had stable richness, and decreased 0.7× in diversity of order 1). The stability of the MMS+ bacterial communities was also supported by the β-diversity.

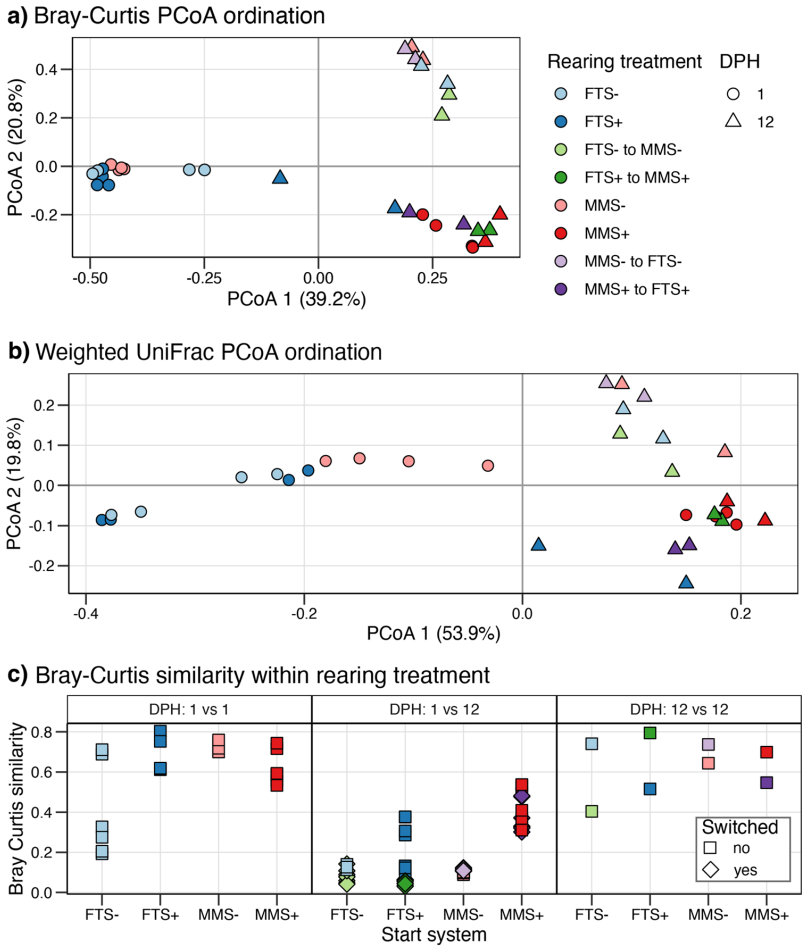


Figure 3. Community composition comparisons between samples (β-diversity) based on rearing treatment. PCoA ordinations are based on (**a**) Bray–Curtis or (**b**) weighted UniFrac distance. Colours indicate rearing treatment, and shape signifies 1 (filled circle) and 12 DPH (filled triangle). (**c**) The Bray–Curtis similarity within rearing treatment within and between sampling days. Colours indicate rearing condition and shape unswitched (filled square) and switched (filled diamond) treatments.

The MMS+ rearing bacterial community was most stable over time. The differences in bacterial community composition between samples were quantified using Bray–Curtis and the weighted UniFrac distances and then ordinated using PCoA (Fig. 3). The PCoA ordinations indicated that most of the differences in community composition were explained by sampling day and rearing treatment (Fig. 3a,b). The MMS+ samples clustered oppositely from the other three rearing treatments along Axis 1 at 1 DPH. Axis 1 explained 39.2% (Bray–Curtis) and 53.9% (UniFrac) of the variation in the distance matrixes, indicating that there was a large difference in community composition between MMS+ and the other treatments. At 1 DPH, the FTS+, FTS− and MMS− clustered together in the Bray–Curtis ordination but were more spread out when using the weighted UniFrac distance. As UniFrac is based on phylogenetic community dissimilarity, this spreading indicates that the ASVs that contributed to community differences between treatments were more different phylogenetically.

At 12 DPH, the differences in the bacterial community composition were separated based on the microbial carrying capacity along Axis 2. This axis explained 20.8% (Bray–Curtis) and 19.8% (UniFrac) of the variation. Moreover, we observed that all 12 DPH samples clustered closer to the 1 DPH MMS+ samples regardless of rearing treatment. This pattern indicated that succession drove the communities toward a common bacterial community composition. The MMS+ samples had already obtained this composition at 1 DPH, highlighting the advantage of pre-feeding the biofilter to acquire a stable microbial community composition.

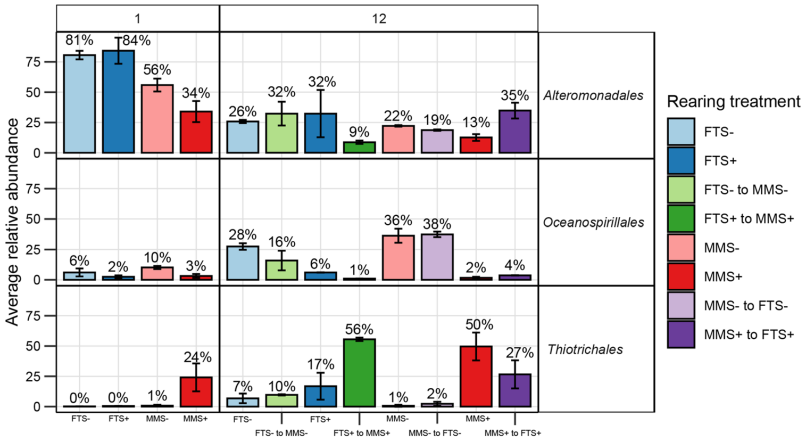


Figure 4. The relative abundance of the three most dominating orders in the dataset. These orders had a > 20% abundance in a minimum of two samples. Colours indicate the rearing treatment. The average relative abundance is shown on each sampling day, and whiskers represent the standard deviation.

The stability of the bacterial community composition was investigated by quantifying the within-system Bray–Curtis similarity within and between sampling days (Fig. 3c). The tanks starting with the MMS+ treatment had the highest bacterial community similarity when comparing 1 and 12 DPH with an average Bray–Curtis similarity of 0.4. In comparison, the Bray Curtis similarity was, on average, 0.1 in tanks starting with the other treatments (Kruskal–Wallis $p < 0.001$).

Next, we evaluated if the initial rearing condition had left a legacy effect on community composition. We compared the Bray–Curtis similarity at 12 DPH between switched and unswitched communities. Unfortunately, we could not perform statistics on these observations due to low power within the groups. The 'FTS– to MMS–' bacterial communities had an average Bray–Curtis similarity (± SD) of 0.4 (± 0.08) and 0.6 (± 0.05) to the communities of the MMS– and FTS–, respectively. The 'MMS– to FTS–' samples showed a similar pattern, with slightly higher similarity to communities continuing with the same initial treatment with average Bray–Curtis similarities of 0.6 (± 0.03) and 0.5 (± 0.01) to the MMS- and FTS- treatments, respectively. Thus, some legacy effects on the bacterial community composition might have established in both the MMS- and the FTS- tanks, but these effects were minor. Clearer patterns were observed in the conditions with high carrying capacity.

The bacterial communities switching from 'MMS+ to FTS+' resisted a change toward the FTS+ community structure. Instead, these 'MMS+ to FTS+' communities had higher Bray–Curtis similarities to the communities continuing with the MMS+ treatment (0.5 ± 0.1) than tanks that initially got the FTS+ treatment (0.2 ± 0.05). This is an indication of a legacy effect in the MMS+ rearing tanks. However, we observed the opposite for the 'FTS+ to MMS+' communities, which had higher Bray–Curtis similarity to the MMS+ communities (0.7 ± 0.06) than those continuing with FTS+ (0.4 ± 0.07). Thus, there was no legacy effect in the FTS+ rearing tanks. Due to the inconsistent patterns, we conclude that the initial rearing condition does not leave a legacy effect on the bacterial community composition. Instead, the mature biofilter (MMS+) supplied a bacterial community that was able to establish quickly in the tanks that previously were FTS+. To evaluate if the MMS+ biofilter seeded a bacterial community, we investigated the taxonomic composition of the samples.

The bacterial community composition in the MMS+ rearing tanks differed taxonomically from those of the other treatments. The class *Gammaproteobacteria* dominated the rearing water in all treatments with an average relative abundance of 76 (± 11% SD). At the order level, we observed differences based on sampling day and rearing treatment (Fig. 4). At 1 DPH, the FTS–, FTS+ and MMS– were similar in bacterial composition, with a high abundance of *Alteromonadales*. The composition was different in the MMS+ rearing water, with substantially lower abundances of *Alteromonadales* and high abundances of *Thiotrichales*. At 12 DPH, the abundance of *Thiotrichales* had doubled in the MMS+ treatment from an average of 24% to 50%. Interestingly, this order also increased in the rearing tanks that switched from 'FTS+ to MMS+'. Its abundance was 56% in the 'FTS+ to MMS+' tanks but only 17% in the FTS+ tanks. This noteworthy difference in abundance indicated that the biofilter community was effectively seeded to the rearing tanks. Next, we investigated if the rearing treatments affected larval viability.

The present rearing treatment had the largest effect on larval performance. Comparing the larval dry weight between the treatments at each sampling day did not indicate that the rearing conditions affected the growth (Supplementary Fig. 2). At 17 and 18 DPH, there was no statistically significant difference between the average weight in the different rearing treatments. However, differences were observed in larval robustness.

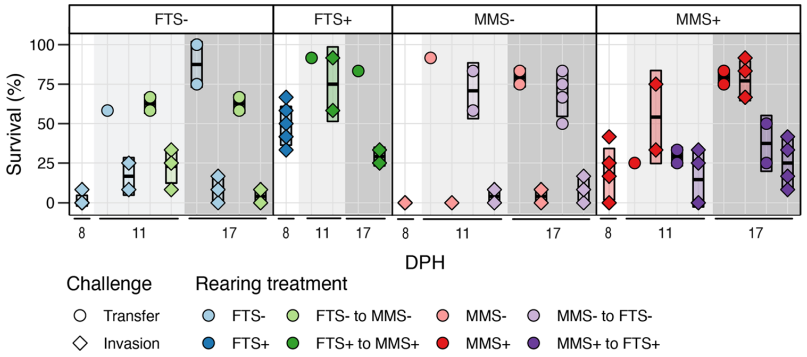


Figure 5. Percent of surviving larvae one day after the transfer and invasion challenge tests at various DPH. Samples are organized based on the initial rearing treatment. Colours indicate the overall rearing treatment. Boxplots represent mean survival ± SD for each rearing treatment at each sampling day.

The robustness of the larvae was investigated in side experiments on 8, 11 and 17 DPH by inducing stress through transfer or exposing the larvae to rearing water invaded with a *Pseudoalteromonas* and a *Polaribacter* bacterial strain (Fig. 5). While *Polaribacter* has been identified as a commensal[16,45], *Pseudoalteronomonas* contains many pathogenic strains towards Atlantic cod[46]. The invaded rearing water thus pose a threat both through an increased bacterial load and exposure to a potentially pathogenic bacterium. Larval mortality was recorded 24 h after the challenge. Not surprisingly, the survival was higher for the larvae only challenged by transfer (mean 68.1 ± 21.2%) compared to larvae transferred to invaded rearing water (mean 20.5 ± 24.8%).

For the larvae that only were subjected to the transfer challenge, differences were observed between the rearing treatments. On average, the survival of larvae was comparable between the FTS−, FTS+ and MMS− treatments but was 1.5× lower for the MMS+ (Fig. 5). Generally, there was no indication that the initial rearing condition affected the general stress level of the larvae. Instead, robustness appeared to be related to the present rearing regime. For example, on 17 DPH, the larvae that continued with MMS+ had 2.1× higher survival than those that switched to FTS+ (i.e. 'MMS+ to FTS+'). Thus, the initial rearing condition left no legacy effect on the general stress level of the fish.

For the invasion challenge, the larvae from tanks with low carrying capacity were the least robust. For these larvae, the mean (± SD) survival was 6.3 (± 8.6)%, and some flasks had 0% survival. In comparison, the larvae from tanks with high carrying capacity had a mean survival of 39.4 (± 26.8)% after invasion stress (Fig. 5). The data from the challenge tests did not indicate that the initial rearing condition left legacy effects on the larval robustness. For example, larvae from tanks that continued in MMS+ challenged with invasion had high survival [mean 69.4 (± 20.2)%], whereas larvae from the tanks that switched from 'MMS+ to FTS+' had 3.5× lower survival [mean 19.8 (± 16.0)%]. Unfortunately, we do not have samples from the FTS+ rearing treatment after 8 DPH due to high mortality in the rearing tanks. If there was a legacy effect, one would expect improved robustness to invasion when switching to a rearing regime associated with higher survival. Furthermore, the larval survival after the challenges was comparable between the FTS− and 'FTS− to MMS−'and between the MMS− and 'MMS− to FTS−'. In conclusion, there was no indication of a legacy effect in the larvae. Instead, the post-switch rearing treatment had the largest impact.

Larval survival was very low in FTS+ tanks. Larval survival at the end of the experiment was comparable and relatively high for the MMS+, MMS− and FTS− treatments. In these treatments, the survival ranged between 12 and 26%. However, survival was low for all tanks that at some point received FTS+ water, ranging from 0 to 7% (Fig. 6). It should be noted that the water quality was visually poorer in the FTS+ tank water. Nevertheless, we investigated if any ASVs were linked to survival.

We identified ASVs with significant log-fold changes between the bacterial communities in high and low survival tanks using a DeSeq2 analysis. Fifty-two ASVs had higher abundances in the communities from tanks with low survival, and 85 had higher abundances in those with high survival (FDR adjusted p-value < 0.05, Supplementary Fig. 3). An interesting pattern emerged when investigating the abundance of the identified ASVs in each rearing tank (Supplementary Fig. 4). At 1 DPH, the abundance of ASVs associated with low survival was over 40% in FTS+, FTS− and MMS− but below 20% in the MMS+ tanks.

When comparing switched and unswitched treatments at 12 DPH, it was apparent that the abundances of these low survival-associated ASVs were treatment dependent. For example, the abundances of these ASVs were 55% in the FTS+ treatment but 3.7× lower in the 'FTS+ to MMS+' treatment. The opposite was observed between MMS+ and 'MMS+ to FTS+'. The low survival associated ASVs were only present at 1% in the MMS+ rearing tanks but increased to 15% in the 'MMS+ to FTS+' tanks. Furthermore, we found five ASVs classified as *Moritella* to be especially interesting. These five ASVs all had over a 7.5-log2 fold increase in the low survival tanks. Four of these ASVs were most similar to the type strain *Moritella viscosa*, a known fish pathogen (Supplementary

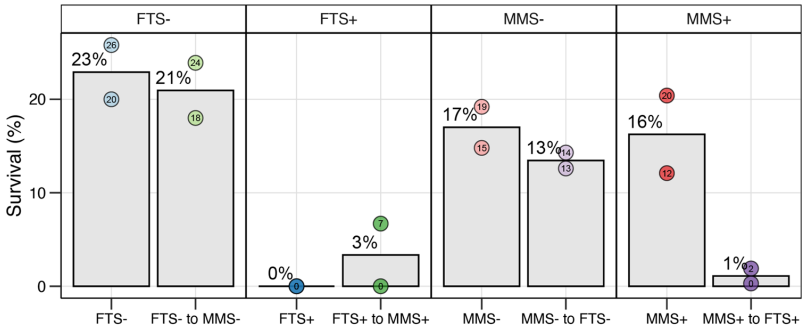


Figure 6. The survival in each rearing treatment at the end of the experiment at 20 DPH. The grey bars and percentages indicate the mean survival in the rearing tanks, whereas the points show each tank's survival.

Table 3, similarity > 92%). Our findings show that the rearing conditions can be used to select for a beneficial microbial environment for the larvae.

Discussion

Marine larvae are hard to rear due to their vulnerability and poor viability associated with high mortality. However, microbial management of the rearing conditions, such as disinfecting the eggs[20] or microbially maturing the water before it reaches the fish[14], has resulted in considerable improvements in larval performance. These improvements indicate that microbial water quality is a major cause of poor larval viability.

We have previously hypothesised that the initial microbial rearing conditions leave legacy effects in Atlantic cod larvae[9], but this had not been tested. Thus, in this paper, we investigated if legacy effects could be observed after initial rearing in either MMS or FTS water treatment in a unique and novel experimental design. This experiment's unique feature was switching the incoming water treatment system in half of the rearing tanks at 9 DPH. This design had two major advantages compared to traditional designs. Firstly, only the inlet system pipes were changed at the water treatment switch. Thus, the larvae were never exposed to the stress of transfer between tanks, allowing us to measure the actual effect of the input microbiota. Secondly, as half of the tanks continued with the same rearing treatment, we could evaluate if legacy effects were established by investigating the larval performance and the microbial community characteristics. Overall, there was no evidence that the initial rearing conditions left a legacy effect in the larvae or their surrounding microbial communities after switching to novel conditions at 9 DPH.

Comparing switched and unswitched rearing regimes did not indicate a legacy effect in the bacterial community of the rearing water. Instead, we found that the microbial carrying capacity and the post-switch water treatment system were the main determinants for differences in bacterial density, growth potential, α-diversity and community composition. However, we observed that the bacterial community of the rearing water that switched from 'MMS+ to FTS+' had higher Bray–Curtis similarity to the community of the rearing water in MMS+ than FTS+ water. This observation could indicate a legacy effect in the MMS+ tanks. However, given that no other analysis pointed to a legacy effect, it instead seems like these observations reflect the stability of the MMS+ biofilter biofilm and the dominant effect of this biofilm on the rearing water. It is likely that this biofilm supplied, or seeded, a stable flow of inlet bacteria to the rearing water. This biofilm seeding effect was highest in the MMS+ treatment. The water provided by the MMS+ had a 10× higher bacterial density than the water from the MMS-. This difference in bacterial load might explain why the seeding effect was less pronounced in the MMS- rearing tanks. From these results, we recommend using microbially matured water at a carrying capacity similar to the carrying capacity in the rearing water to obtain a stable microbial rearing environment. This is in accordance with previous studies[7].

Compared to the rearing water, the fish as a microbial ecosystem is less affected by the high water exchange rates. As such, we wanted to see if a legacy effect established in the larvae. Neither survival, weight, nor robustness of the larvae indicated that the larvae experienced a legacy effect based on the initial rearing conditions. Theoretically, for a legacy effect to establish, the effect needs to manifest in a deterministic way. Fish as an ecosystem is not stable at 20 DPH[47]. Due to large morphological and physiological changes during larval development, the niches available on and in the fish will change[48,49]. The lack of a legacy effect in Atlantic cod larvae is consistent with legacy-effects studies in other fish species, such as Nile tilapia (*Oreochromis niloticus*)[50] and zebrafish (*Danio rerio*)[49]. This consistency poses the question of whether one can detect legacy effects in fish larvae that, through development, change the deterministic constraints of their environment. It has been documented that both deterministic and stochastic processes structure the fish larvae's microbiome [16,51]. However, it is unclear what drives the deterministic processes, if these drivers are stable over time, and if the initial environmental conditions can impact them. In this experiment, we could not find that the initial rearing condition established legacy effects in the system. However, other drivers might be affected by historical effects. Our findings suggest that the fish developmental stage and environmental microbiota have most impact on the fish microbiota composition.

Instead of a legacy effect, the present rearing condition appeared important for larval survival and robustness. Especially pronounced was the low larval survival in tanks that at some timepoint were connected to the FTS+ rearing treatment (range 0–7%). In a previous experiment, the survival of Atlantic cod larvae was 65% higher in MMS compared to FTS systems at 32 DPH[9]. In that experiment, the rearing water system did not impact the bacterial community stability. The authors hypothesised that the higher survival in MMS was linked to a beneficial initial microbial colonisation and earlier onset of growth of the larvae upon mouth opening. Whilst we did observe a major difference in survival between larvae reared in MMS+ and FTS+, no difference was observed between MMS- and FTS-. Most importantly, there was no difference in survival between tanks reared entirely in FTS+ and those that switched from 'MMS+ to FTS+'. Thus, the initial protective colonization from an assumed more beneficial microbiome (MMS) did not remain in the larvae.

This protective effect was also lacking in terms of larval robustness. Rather than a legacy effect, the largest differences in larval robustness were based on the microbial carrying capacity of the tanks. We observed that larvae from tanks with high carrying capacity were more robust to invasion by bacteria than those reared in low carrying capacity. This difference is likely related to the propagule pressure or the relative abundance of the invader. Higher propagule pressures increase the probability of invasion success[52]. In the high carrying capacity tanks, the bacterial density was, on average, 7.8× higher. Thus, these bacterial communities experienced a significantly lower propagule pressure when invaded. As marine larvae drink approximately 10^4–10^6 bacteria from the rearing water per day[10,53], the larvae in the high carrying capacity water had a lower chance of being colonised or exposed to the invading bacteria. This observation might explain why higher survival and robustness are observed in high carrying capacity aquaculture systems such as recirculating aquaculture systems (RAS) and fed MMS systems[9,12,13,54].

An unexpected observation was that the larvae from the FTS+ tanks were most robust to challenges but had the lowest survival at the end of the experiment. A likely explanation for this contrasting behaviour is that the larvae with low fitness already had died in the tanks. Consequently, when sampling from the tanks with few individuals left, the likelihood of selecting more robust larvae was higher in the FTS+ tanks. Thus, these results are likely biased towards a difference in population fitness means. The differences were caused by environmental conditions rather than a legacy effect.

Instead of a legacy effect from the initial rearing condition, microbial selection within the rearing tanks seemed to contribute to larval performance. Our rearing water investigations indicate that we were able to provide a continuous input of a stable microbiota from the MMS+ biofilter. Communities with high microbial community stability are usually found in K-selective environments[12,54]. Such environments typically have a stronger selection pressure and potentially a higher utilisation of available niches. In contrast, the three other treatments (MMS−, FTS− and FTS+) were prone to higher microbial turnover. Previous studies have documented that a large gap between the carrying capacity of the incoming and rearing water select for fast-growing opportunistic r-strategic bacteria[3]. Therefore, good rearing management should avoid r-selective environments, as most detrimental bacteria are r-strategists. We observed that excess available resources in the rearing tanks resulted in an enormous increase in bacterial density (132× in FTS+), indicating a strong r-selective environment. On the other hand, the MMS+ that received the same amount of resources, but added in the biofilter before the tanks, appeared to be K-selective, as the bacterial density was stable between the incoming and rearing water (1.2×). Thus, we obtained different selective environments in our treatments.

That the FTS+ rearing water was r-selected was also reflected by the ASVs linked to low survival. Although a difference in oxygen saturation might have contributed to mortality, 5 of the 52 ASVs linked to low survival had high 16S rRNA gene similarity to the type strain of *Moritella viscosa*. This bacterium is known to cause winter ulcers in cold-water fish such as Atlantic salmon and Atlantic cod[55]. Its increased abundance in the tanks with low survival might have contributed to the higher mortality in these tanks. For example, increased pathogen concentration has been shown to activate the adaptive immune systems[16], making the fish more stressed. It should be noted that we did not study the larval microbiota. Vestrum et al. (2020) showed that an OTU belonging to the *Arcobacter* genus dominated the cod larval microbiota with abundances over 65%, whereas in the rearing water it never exceeded 2.3%[16]. Thus, investigations into the larval microbiome in this experiment might have allowed us to identify a potential pathogen with higher certainty.

The differences in survival between the treatments have two major implications. First, a protective legacy effect was not observed in the larvae. Instead, it seems like the microbial selection pressure in the rearing tanks provided a protective effect. The probability that a larva encounters a detrimental bacterium in K-selective environments is lower than in r-selected environments. This was apparent when comparing the high survival in tanks with the MMS+ treatment throughout the experiment with the very low survival in tanks that switched to FTS+. Thus, exposure to unfavourable conditions increases mortality regardless of how good the conditions are during the initial hatching. Secondly, when comparing the survival in the FTS+ and 'FTS+ to MMS+' tanks, there was no indication that improving the rearing environment led to higher survival. This indicates that if the larvae are exposed to unfavourable conditions early in life, their viability cannot be reversed by simply improving their environment. These observations have significant implications and illustrate that to obtain high survival and larval welfare it is vital to have good rearing conditions from the start and throughout the larval rearing period.

In conclusion, we found no evidence that the initial rearing condition left legacy effects on Atlantic cod larvae or their rearing water microbiota. Instead, the difference in carrying capacity between the intake and rearing water, the rearing water carrying capacity and the present water treatment had a much higher importance for larval viability and microbiota characteristics. We are the first to investigate legacy effects and report the lack of these effects during the first twenty days of marine Atlantic cod larvae rearing. Our study emphasises the importance of providing a beneficial rearing microbial environment throughout the rearing period to obtain high larval viability and bacterial water quality.

Data availability

The sequencing data is available at ENA under study ERP138000 with accession numbers ERR9837055-ERR9837086. In addition, raw data for bacterial density, bacterial net growth potential, larval weight and larval survival (challenge and final) are available as supplementary files. All scripts used to perform data analysis and plot generation are available at https://github.com/madeleine-gundersen/legacy_effects_in_rearing_systems.

Received: 20 June 2022; Accepted: 9 November 2022
Published online: 17 November 2022

References

1. Llewellyn, M. S., Boutin, S., Hoseinifar, S. H. & Derome, N. Teleost microbiomes: The state of the art in their characterization, manipulation and importance in aquaculture and fisheries. *Front. Microbiol.* **5**, 207. https://doi.org/10.3389/fmicb.2014.00207 (2014).
2. Yukgehnaish, K. *et al.* Gut microbiota metagenomics in aquaculture: Factors influencing gut microbiome and its physiological role in fish. *Rev. Aquacult.* **12**, 1903–1927. https://doi.org/10.1111/RAQ.12416 (2020).
3. Vadstein, O., Attramadal, K. J. K., Bakke, I. & Olsen, Y. K-Selection as microbial community management strategy: A method for improved viability of larvae in aquaculture. *Front. Microbiol.* **9**, 2730. https://doi.org/10.3389/fmicb.2018.02730 (2018).
4. Blancheton, J. P., Attramadal, K. J. K., Michaud, L., d'Orbcastel, E. R. & Vadstein, O. Insight into bacterial population in aquaculture systems and its implication. *Aquacult. Eng.* **53**, 30–39. https://doi.org/10.1016/J.AQUAENG.2012.11.009 (2013).
5. Håstein, T. Animal welfare issues relating to aquaculture. in *Global Conference on Animal Welfare: An OIE Initiative*. 219–231. (Office for Official Publications of the European Community, 2004).
6. Brauner, C. J. & Richards, J. G. Physiological performance in aquaculture: Using physiology to help define optimal conditions for growth and environmental tolerance. *Fish Physiol. (Academic Press)* **38**, 83–121. https://doi.org/10.1016/bs.fp.2020.10.001 (2020).
7. de Schryver, P. & Vadstein, O. Ecological theory as a foundation to control pathogenic invasion in aquaculture. *ISME J.* **8**, 2360–2368. https://doi.org/10.1038/ismej.2014.84 (2014).
8. Borges, N. *et al.* Bacteriome structure, function, and probiotics in fish Larviculture: The good, the bad, and the gaps. *Annu. Rev. Anim. Biosci.* **9**, 423–452. https://doi.org/10.1146/annurev-animal-062920-113114 (2021).
9. Attramadal, K. J. K. *et al.* RAS and microbial maturation as tools for K-selection of microbial communities improve survival in cod larvae. *Aquaculture* **432**, 483–490. https://doi.org/10.1016/j.aquaculture.2014.05.052 (2014).
10. Vadstein, O. *et al.* Managing the microbial community of marine fish larvae: A holistic perspective for larviculture. *Front. Microbiol.* **9**, 1820. https://doi.org/10.3389/fmicb.2018.01820 (2018).
11. Attramadal, K. J. K. *et al.* Microbial maturation of intake water at different carrying capacities affects microbial control in rearing tanks for marine fish larvae. *Aquaculture* **457**, 68–72. https://doi.org/10.1016/j.aquaculture.2016.02.015 (2016).
12. Attramadal, K. J. K. *et al.* Recirculation as a possible microbial control strategy in the production of marine larvae. *Aquacult. Eng.* **46**, 27–39. https://doi.org/10.1016/j.aquaeng.2011.10.003 (2012).
13. Attramadal, K. J. K. *et al.* The effects of moderate ozonation or high intensity UV-irradiation on the microbial environment in RAS for marine larvae. *Aquaculture* **330–333**, 121–129. https://doi.org/10.1016/j.aquaculture.2011.11.042 (2012).
14. Skjermo, J., Salvesen, I., Øie, G., Olsen, Y. & Vadstein, O. Microbially matured water: A technique for selection of a non-opportunistic bacterial flora in water that may improve performance of marine larvae. *Aquacult. Int.* **5**, 13–28. https://doi.org/10.1007/BF02764784 (1997).
15. Vellend, M. *The Theory of Ecological Communities*. Vol. 57 (Princeton University Press, 2016).
16. Vestrum, R. I., Attramadal, K. J. K., Vadstein, O., Gundersen, M. S. & Bakke, I. Bacterial community assembly in Atlantic cod larvae (*Gadus morhua*): Contributions of ecological processes and metacommunity structure. *FEMS Microbiol. Ecol.* **96**, 163. https://doi.org/10.1093/femsec/fiaa163 (2020).
17. Vass, M. & Langenheder, S. The legacy of the past: Effects of historical processes on microbial metacommunities. *Aquatic Microb. Ecol.* **79**, 13–19. https://doi.org/10.3354/ame01816 (2017).
18. Symons, C. C., Schulhof, M. A., Cavalheri, H. B. & Shurin, J. B. Legacy effects of fish but not elevation influence lake ecosystem response to environmental change. *J. Anim. Ecol.* **90**, 662–672. https://doi.org/10.1111/1365-2656.13398 (2021).
19. Cuddington, K. Legacy effects: The persistent impact of ecological interactions. *Biol. Theory* **6**, 203–210. https://doi.org/10.1007/S13752-012-0027-5 (2012).
20. Salvesen, I., Øie, G. & Vadstein, O. Surface disinfection of Atlantic halibut and turbot eggs with glutaraldehyde: Evaluation of concentrations and contact times. *Aquacult. Int.* **5**, 249–258. https://doi.org/10.1023/A:1018343602872 (1997).
21. Ministry of Agriculture and Food & Norway. *Animal Welfare Act*. https://www.regjeringen.no/en/dokumenter/animal-welfare-act/id571188/ (2009).
22. Attramadal, K. J. K. *et al.* Ceramic clay reduces the load of organic matter and bacteria in marine fish larval culture tanks. *Aquacult. Eng.* **49**, 23–34. https://doi.org/10.1016/j.aquaeng.2012.02.003 (2012).
23. Cole, J. R. *et al.* Ribosomal Database Project: Data and tools for high throughput rRNA analysis. *Nucleic Acids Res.* **42**, D633–D642. https://doi.org/10.1093/nar/gkt1244 (2013).
24. Gundersen, M. S., Morelan, I. A., Andersen, T., Bakke, I. & Vadstein, O. The effect of periodic disturbances and carrying capacity on the significance of selection and drift in complex bacterial communities. *ISME Commun.* **1**, 53. https://doi.org/10.1038/s43705-021-00058-4 (2021).
25. *NSC-Sequencing*. https://www.sequencing.uio.no/.
26. Edgar, R. C. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**, 2460–2461. https://doi.org/10.1093/bioinformatics/btq461 (2010).
27. Edgar, R. C. UNOISE2: Improved error-correction for Illumina 16S and ITS amplicon sequencing. *bioRxiv* 81257. https://doi.org/10.1101/081257 (2016).
28. Edgar, R. C. SINTAX: A simple non-Bayesian taxonomy classifier for 16S and ITS sequences. *bioRxiv* 74161. https://doi.org/10.1101/074161 (2016).
29. R Core Team. *R: A Language and Environment for Statistical Computing*. (2020).
30. Wright, E. S. DECIPHER: Harnessing local sequence context to improve protein multiple sequence alignment. *BMC Bioinform.* **16**, 322. https://doi.org/10.1186/s12859-015-0749-z (2015).
31. Schliep, K. P. phangorn: Phylogenetic analysis in R. *Bioinformatics* **27**, 592–593. https://doi.org/10.1093/bioinformatics/btq706 (2011).
32. Paradis, E., Claude, J. & Strimmer, K. APE: Analyses of phylogenetics and evolution in R language. *Bioinformatics* **20**, 289–290. https://doi.org/10.1093/bioinformatics/btg412 (2004).
33. Lahti, L. & Sudarshan, S. *Microbiome: Tools for Microbiome Analysis in R*. (2017).
34. McMurdie, P. J. & Holmes, S. phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS ONE* **8**, e61217. https://doi.org/10.1371/journal.pone.0061217 (2013).

35. Wickham, H. ggplot2. *WIREs Comput. Stat.* **3**, 180–185. https://doi.org/10.1002/wics.147 (2011).
36. Kassambaram, A. *ggpubr: 'ggplot2' Based Publication Ready Plots*. (2020).
37. Lucas, R. *et al.* A critical evaluation of ecological indices for the comparative analysis of microbial communities based on molecular datasets. *FEMS Microbiol. Ecol.* https://doi.org/10.1093/femsec/fiw209 (2017).
38. Dixon, P. VEGAN, a package of R functions for community ecology. *J. Veg. Sci.* **14**, 927–930. https://doi.org/10.1111/j.1654-1103.2003.tb02228.x (2003).
39. Chambers, J. M. & Hastie, T. J. *Statistical Models in S. Philosophy* (Chapman and Hall/CRC, 1992).
40. Conover, W. J., Johnson, M. E. & Johnson, M. M. A comparative study of tests for homogeneity of variances, with applications to the outer continental shelf bidding data. *Technometrics* **23**, 351–361. https://doi.org/10.1080/00401706.1981.10487680 (1981).
41. Royston, J. P. An extension of Shapiro and Wilk's W test for normality to large samples. *J. R. Stat. Soc. Ser. Appl. Stat.* **31**, 115–124. https://doi.org/10.2307/2347973 (1982).
42. Hollander, M., Wolfe, D. A. & Chicken, E. *Nonparametric Statistical Methods* (Wiley, 2013).
43. Tukey, J. W. Comparing individual means in the analysis of variance. *Biometrics* **5**, 99. https://doi.org/10.2307/3001913 (1949).
44. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550. https://doi.org/10.1186/s13059-014-0550-8 (2014).
45. Vestrum, R. I. *et al.* Commensal and opportunistic bacteria present in the microbiota in Atlantic cod (*Gadus morhua*) larvae differentially alter the hosts' innate immune responses. *Microorganisms*. https://doi.org/10.3390/MICROORGANISMS10010024/S1 (2022).
46. Sandlund, N. & Bergh, Ø. Screening and characterisation of potentially pathogenic bacteria associated with Atlantic cod *Gadus morhua* larvae: Bath challenge trials using a multidish system. *Dis. Aquat. Organ* **81**, 203–217. https://doi.org/10.3354/DAO01934 (2008).
47. Yan, Q. *et al.* Environmental filtering decreases with fish development for the assembly of gut microbiota. *Environ. Microbiol.* **18**, 4739–4754. https://doi.org/10.1111/1462-2920.13365 (2016).
48. Stephens, W. Z. *et al.* The composition of the zebrafish intestinal microbial community varies across development. *ISME J.* **10**, 644–654. https://doi.org/10.1038/ismej.2015.140 (2015).
49. Xiao, F. *et al.* Host development overwhelms environmental dispersal in governing the ecological succession of zebrafish gut microbiota. *NPJ Biofilms Microbiomes* **7**, 5. https://doi.org/10.1038/s41522-020-00176-2 (2021).
50. Deng, Y., Kokou, F., Eding, E. H. & Verdegem, M. C. J. Impact of early-life rearing history on gut microbiome succession and performance of Nile tilapia. *Anim. Microbiome* **3**, 81. https://doi.org/10.1186/s42523-021-00145-w (2021).
51. Burns, A. R. *et al.* Contribution of neutral processes to the assembly of gut microbial communities in the zebrafish over host development. *ISME J.* **10**, 655–664. https://doi.org/10.1038/ismej.2015.142 (2015).
52. Vila, J. C. C., Jones, M. L., Patel, M., Bell, T. & Rosindell, J. Uncovering the rules of microbial community invasions. *Nat. Ecol. Evolut.* **3**, 1162–1171. https://doi.org/10.1038/s41559-019-0952-9 (2019).
53. Reitan, K. I., Natvik, C. M. & Vadstein, O. Drinking rate, uptake of bacteria and microalgae in turbot larvae. *J. Fish Biol.* **53**, 1145–1154. https://doi.org/10.1111/J.1095-8649.1998.TB00238.X (1998).
54. Salvesen, I., Skjermo, J. & Vadstein, O. Growth of turbot (*Scophthalmus maximus* L.) during first feeding in relation to the proportion of r/K-strategists in the bacterial community of the rearing water. *Aquaculture* **175**, 337–350. https://doi.org/10.1016/S0044-8486(99)00110-6 (1999).
55. Colquhoun, D. J., Hovland, H., Hellberg, H., Haug, T. & Nilsen, H. Moritella viscosa isolated from farmed Atlantic cod (*Gadus morhua*). *Bull. Eur. Assoc. Fish Pathol.* **24**, 109–114 (2004).

Acknowledgements

This study was financed by AQUAEXCEL TNA project 0065/04/10/25 and the COFASP ERA-Net project "MicStaTech" funded by the Research Council of Norway (contract no. 247558). Financial support was also provided by NTNU, Faculty of Natural Sciences, as a PhD scholarship to MSG. We want to thank Ian Moreland for performing DNA extraction, Mari-Ann Østensen for her help with this study and Alexander Fiedler for providing valuable feedback on the manuscript.

Author contributions

M.S.G. performed the amplicon sequencing and subsequent data analysis and prepared the first draft of the manuscript. K.A. and P.D. designed and conducted the study, with input from O.V. All authors reviewed and developed the final manuscript.

Funding

Open access funding provided by Norwegian University of Science and Technology.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-022-24149-x.

Correspondence and requests for materials should be addressed to M.S.G.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Paper IV

The early gut microbiome of wild and aquaculture strains of Atlantic salmon is influenced by stochastic processes and environmental bacteria

This paper is not yet published and is therefore not included.

Paper V

The impact of phage treatment on bacterial community structure is minor compared to antibiotics

The impact of phage treatment on bacterial community structure is minor compared to antibiotics

Madeleine S. Gundersen[a], Alexander W. Fiedler[a], Ingrid Bakke[a], Olav Vadstein[a]

[a]Department of Biotechnology and Food Science, Norwegian University of Science and Technology (NTNU), Norway

Corresponding author: Olav Vadstein (olav.vadstein@ntnu.no) & Madeleine S Gundersen (madeleine.gundersen@ntnu.no)

Abstract

Background: Phage treatment is suggested as an alternative to antibiotics; however, there is limited knowledge of how phage treatment impacts resident bacterial community structure. When phages induce bacterial lysis, resources become available to the resident community. Therefore, the density of the target bacterium is essential to consider when investigating the effect of phage treatment. This has never been studied. Thus, we invaded microcosms containing a lake-derived community with *Flavobacterium columnare* strain Fc7 at no, low or high densities, and treated them with either the bacteriophage FCL-2, the antibiotic Penicillin or kept them untreated (3x3 factorial design). The communities were sampled over the course of one week, and bacterial community composition and density were examined by 16S rDNA amplicon sequencing and flow cytometry.

Results: We show that phage treatment had minor impacts on the resident community when the host *F. columnare* Fc7 of the phage was present, as it caused no significant differences in bacterial density α- and β-diversity, successional patterns, and community assembly. However, a significant change was observed in community composition when the phage host was absent, mainly driven by a substantial increase in *Aquirufa*. In contrast, antibiotics induced significant changes in all community characteristics investigated. The most crucial finding was a bloom of *γ-proteobacteria* and a shift from selection to ecological drift dominating community assembly.

Conclusions:

This study investigated whether the amount of a bacterial host impacted the effect of phage treatment on community structure. We conclude that phage treatment did not significantly affect the diversity or composition of the bacterial communities when the phage host was present, but introduced changes when the host was absent. In contrast, antibiotic treatment was highly disturbing to community structure. Moreover, higher amounts of the bacterial host of the phage increased the contribution of stochastic community assembly and resulted in a feast-famine like response in bacterial density in all treatment groups. This finding emphasises that the invader density used in bacterial invasion studies impacts the experimental reproducibility. Overall, this study supports that phage treatment is substantially less disturbing to bacterial communities than antibiotic treatments.

Background

Over the last seven decades, antibiotics have been the dominating method to treat bacterial infections [1]. However, their overuse and misuse have led to antibiotic-resistant bacteria at a fast rate [2], and it is now evident that antibiotics can negatively impact the resident bacterial community [3]. As a result, it is critical to find alternatives to antibiotics to combat bacterial infections [4, 5]. One potential alternative is phage therapy, which utilises viruses that specifically kill target bacteria [6].

Antibiotics have numerous times been documented to affect the resident microbial community, leading to, for example, decreased bacterial growth, diversity, stability, and functionality and causing overall changes in community composition [3, 7–9]. In many reports, antibiotic usage disrupts the community beyond its resilience, leaving lasting effects on the microbiome [7, 10]. Changes in the competitive fitness of the populations [11] and broken interaction networks [9] might be underlying ecological mechanisms for the lasting disturbance effects. Because of the ever-growing evidence for the importance of the microbiome for ecosystem functioning, developing treatments that do not disturb the resident bacterial community is imperative [12, 13].

Phages are vital constituents of natural ecosystems [14, 15] and are critical for sustaining high productivity and ecosystem turnover [16]. The phages have a very narrow host specificity, even down to the strain level, and are self-propagating when the host is present [17]. Their therapeutic potential was explored in 1919 by d'Hérelle, just two years after he discovered bacteriophages, and is known as phage therapy [6].

It is frequently stated that phage therapy does not disturb the resident bacterial community [18, 19]. However, few studies have investigated the impact of phage therapy on

the resident bacterial community, and conclusions vary. Most studies have been conducted in animals and humans without the presence of the bacterial host, and aimed to evaluate the safety of phage consumption [20–28]. Moreover, although phage therapy is proposed as an alternative to antibiotics, hardly any studies have compared the impact of phage- versus antibiotic treatment in communities containing the bacterial host of the phage [29–31]. In these studies, phage treatment induced changes in the bacterial communities of rabbits [29], mice [30] and a human-gut synthetic community [31], but was less disruptive to the bacterial communities than antibiotics.

Furthermore, if phage treatment replaces antibiotics, it is critical to evaluate the impact of phage introduction on environmental ecosystems. To our knowledge, only two studies have addressed this issue, but neither had a relevant disturbance control. When phages were added to water, no significant changes were observed in the bacterial community [32], while significant changes were observed when phages were added to soil [33]. Importantly, no studies have manipulated the density of the phage's bacterial host.

From an ecological perspective, one would expect that the abundance of the host bacterium would influence the effects of the phage treatment on the resident community. This is because the lysis of the target bacterium releases mineral nutrients and dissolved organic material (DOM) that stimulate the growth of the resident community [34]. Overall, the impact of phage therapy on the resident microbiome vary and may depend on the population size of the host, and the mechanisms behind the effects are unclear.

In this study, we investigated if phage treatment and different amount of phage-host impacted bacterial community structure. We hypothesised that higher amounts of the bacterial host would release more resources due to bacterial lysis, leading to a greater impact of the treatments on the community composition. A bacterial community from the planktonic

fraction of a natural lake ecosystem was added the bacterial host of the phage at two different densities or left uninvaded, and the added the phage. We used the broad-spectrum antibiotic Penicillin as a positive control for bacterial community disturbance. Both phage and antibiotic treatment cause bacterial lysis. We investigated the impact of phage and antibiotic treatment on community cell density, structure, diversity, and assembly and to evaluated if the amount of phage host modulates these effects. We analysed the bacterial community over a week using 16S rRNA gene amplicon sequencing and flow cytometry.

Methods

Experimental design

A 3x3 factorial experiment was conducted to compare the effect of phage- and antibiotic treatment on resident bacterial community properties after an invasion by *Flavobacterium columnare* strain Fc7 (**Figure 1**). We varied the amount of the invader (no, low and high) and the treatment type (phage, antibiotic or none) to obtain nine experimental groups. Each experimental group comprised five replicates yielding a total of 45 microcosms (250mL cell culture flasks with ventilated caps). The experiment was conducted at 14 °C. The microcosms were invaded at day 0, and treatments were applied one hour after the invasion. Phage- and antibiotic treatment was applied once by adding phage FCL-2 to a multiplicity of infection (MOI) of 2.3 (low) and 2.9 (high) or adding 1mg/L of the antibiotic Penicillin. To secure community turnover, we exchanged 11% of the microcosm daily with 0.2µm-filtered, autoclaved lake water, keeping the volume constant at 100mL. The experiment was terminated after 7 days.

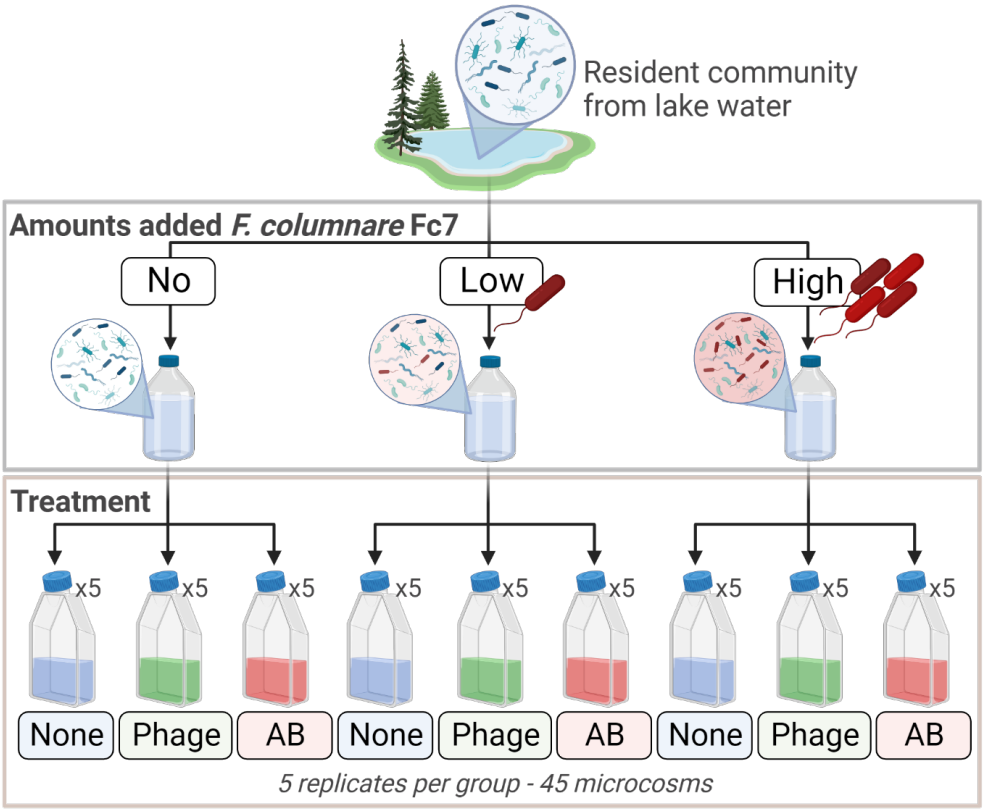


Figure 1: *Flavobacterium columnare* strain Fc7 was added to a lake bacterial community (i.e. the resident community) in three different amounts (no, low or high addition). After mixing the resident community and *F. columnare* Fc7, the communities were split into five experimental microcosms (100 mL each) and received treatment. The communities were treated with nothing ("none"), the phage FCL-2 (Phage) or the antibiotic Penicillin (AB). The experiment consisted of nine groups replicated five times resulting in a total of 45 microcosms.

***Flavobacterium columnare* strain Fc7**

The Gram-negative freshwater bacterium *Flavobacterium columnare* strain Fc7 was used as the invader [35]. *F. columnare* Fc7 was cultivated in TYES medium at room temperature (22°C) under aerobic conditions and with shaking [35]. *F. columnare* Fc7 was taken from a glycerol stock three days before starting the experiment, and 5% v/v was transferred daily in liquid TYES to keep the culture in the exponential phase. One mL of an *F. columnare* Fc7 culture in the late exponential phase was harvested by centrifugation (13,000g, 5min) and resuspended

in 1mL 0.2µm filtered lake water. The density of the harvested *F. columnare* Fc7 was quantified to be $3.4x10^8$ cells/mL using flow cytometry.

Resident bacterial community

The resident bacterial community was collected at 50m depth from a lake (Jonsvatnet, Trondheim, Norway) in May 2022. The collected water was filtered with a 55µm screen to remove larger protozoa and had a bacterial density of $6.25x10^5$cells/mL. The filtered lake water was split into three 2L bottles. *F. columnare* Fc7 was introduced to two of these bottles by adding low ($1.38x10^5$ *F. columnare* Fc7 cells/mL. 24% bacterial density increase) or high amounts of *F. columnare* Fc7 ($1.07x10^6$ *F. columnare* Fc7 cells/mL, 190% bacterial density increase=high) at day 0. The bottles were shaken well, and 100mL was transferred to each microcosm before the treatments were applied.

Phage treatment with FCL-2

The phage FCL-2, which targets *F. columnare* Fc7, was used for the phage treatment [36]. Selectivity towards *F. columnare* Fc7 was confirmed using the soft-agar overlay technique and spot testing [37]. We prepared an FCL-2 phage stock containing 10^{10}PFU/mL. More details on phage stock preparation are given in Supplementary methods. We added $3.12x10^5$PFU/mL to the microcosms with low amounts of *F. columnare* Fc7 added and $3.12x10^6$PFU/mL to the microcosms added high amounts of *F. columnare* Fc7 to obtain an MOI of 2.3 and 2.9, respectively. $3.12x10^6$PFU/mL were also added to the microcosms without *F. columnare* Fc7 added to account for the impact of phages when no host is present.

Estimation of total and living bacterial community density

Each day before the water exchange, 1mL from each microcosm was fixed with glutaraldehyde (0.1% final), snap-frozen and stored at -80°C (45 samples/day). We sampled

an additional 1mL from replicate microcosm #3 for immediate live-dead cell density analysis (9 samples/day). The bacterial density was quantified using flow cytometry (Attune NxT, ThermoFisher). Fixed samples were stained with the RNA-binding fluorescent stain SYBR green II (Invitrogen) to quantify the total bacterial density. To quantify the living population, we immediately stained samples with two dyes; the fluorescent DNA-binding dyes SYBR green I (Invitrogen) and propidium iodide (PI), which enter all or only membrane-compromised cells, respectively. An in-depth description of the protocol, system configurations and gating strategy is given in the Supplementary methods.

Sampling for bacterial community characterisation

At day 1, 3 and 7, 10mL of the water from each microcosm was filtered through a 0.2µm polycarbonate filter (Osmonics, 25mm) to sample the bacterial community. In addition, two technical replicates were sampled from the original lake water before the invasion and two technical replicates from each 2L flask with varying invader density (3 groups). The filters were placed in 1.5mL cryo tubes, snap-frozen and stored at -80°C until DNA extraction.

DNA extraction

For extraction of bacterial DNA, each filter was cut into pieces and homogenised in 750µL DNA/RNA Shield solution (Zymo Research) in ZR BashingBead Lysis Tubes (0.1- and 0.5-mm matrix) using a Precellys 24 (5500rpm-2x30s-15s break, Bertin Technologies). Next, DNA was extracted and purified using the ZymoBiomics MagBead DNA/RNA kit (R2135, Zymo Research) and the KingFisher Flex automated extraction instrument according to the manufacturer's protocol, except for eluting DNA in 100µL water instead of 50µL. Extracted DNA was stored at -20°C.

16S rRNA gene amplicon sequencing and processing

The V3-V4 region of the 16S rRNA gene was amplified using the broad-coverage PCR primers Ill341F-KL (5´-TCG-TCG-GCA-GCG-TCA-GAT-GTG-TAT-AAG-AGA-CAG-NNN-NCC-TAC-GGG-N-3') and Ill805R (5´-GTC-TCG-TGG-GCT-CGG-AGA-TGT-GTA-TAA-GAG-ACA-GNN-NNG-ACT-CAN-VGG-GTA-TCT-AAK-CC-3'). Each reaction was run for 36 cycles (98°C 15 s, 55°C 20 s, 72°C 20 s) with final concentrations of 0.15µM of each primer, 0.25mM of each dNTP, 1x Phusion buffer HF, 0.015units/µL of Phusion Hot Start II DNA polymerase and 1µL of DNA extracts as the template in 25µL reaction volume. PCR products were examined using electrophoresis on 1% agarose gels (1h, 110V) containing 50µM GelRed (Biotium). The amplicon library was prepared as described previously [38] by first purifying and normalising the amplicons using the SequalPrepTM Normalization Plate Kit (Invitrogen) before samples were dual-indexed with Illumina adapters using PCR (FC-131-2001 and 2003, 10 cycles). The indexed amplicons were normalised and purified again before the library was concentrated using Amicon® Ultra-0.5 Centrifugal Filter Devices. The library was sequenced using MiSeq v3 Illumina sequencing (Illumina, San Diego, CA) employing 300 base pair paired reads at the Norwegian Sequencing Centre at the University of Oslo, Norway. The Illumina sequencing reads were processed using the USEARCH pipeline [39] (v.11). An amplicon sequence variant (ASV) table was generated as described previously [40], except for removing reads<370 base pairs instead of 400.

Statistical analysis

All data analysis was performed in R [41] (v. 4.2.2.) with 3003 as a seed. Quality assurance of the amplicon data and normalising strategy can be found in the Supplementary methods.
All R-scripts are available at github.com/madeleine-gundersen/Phage_impact_community.

α-diversity was investigated as Hill diversity of order 0 (richness), 1 and 2 [42] using the normalised amplicon library. For bacterial density and richness, we fitted a third- and

second-degree, respectively, polynomial mixed model using the $\log_{10}$ transformed density or ASV richness as the response variable and day, treatment, amount of *F. columnare* Fc7 added and the interactions between these as the explanatory variables. Sampling day was added as a random intercept term for each sampling unit. Statistical significance was evaluated by performing a Dunnett test on the estimated marginal means ratio difference between control and either phage- or antibiotic treatment at each sampling day, amount of *F. columnare* Fc7 added and treatment comparison.

The β-diversity was evaluated using the Bray-Curtis and Sørensen similarity. The similarity matrixes were obtained by taking the average of 100 similarity matrixes generated by random subsampling of the ASV-table to 26448 reads [43, 44]. The average Bray-Curtis and average Sørensen similarity matrixes were ordinated using principal coordinate analysis (PcoA). Statistical significance was evaluated with the mean of 100 permutational analysis of variances (PERMANOVA, 999 permutations).

Differential abundance analysis was performed on communities sampled at day 7 using corncob [45], DESeq2 [46] and ANCOMBC [47] using the absolute abundance. As input to all tree methods, we filtered out ASVs from the full dataset with a prevalence and total absolute abundance below 5% and 2500 ASVs/mL, respectively. Because different tools can identify different taxa as significant [48], we conservatively defined ASVs identified by all three methods as having significantly different absolute abundances between the treatments.

Community assembly was investigated by quantifying the change in the similarity between communities of replicate microcosms per day (i.e. replicate similarity rate) [40]. The replicate similarity rate was determined as follows. First, the Bray-Curtis and Sørensen similarity was quantified for each pair of replicate microcosms on each sampling day. Next, we performed mixed linear regressions with similarity as the response variable and DPI,

treatment type, amount of *F. columnare* Fc7 added and the interaction between these as the explanatory variables. Repeated sampling from the same unit was accounted for in the random effects term. The temporal slope rate was interpreted as a replicate similarity rate of change. Positive rates indicate that the community composition between two replicates became more similar over time, indicative of selection dominating community assembly. Negative rates, on the other hand, reflect that the replicates became less similar over time, indicating that drift is dominating community assembly. This similarity rate of change was estimated for each experimental group.

Results

In this study, *Flavobacterium columnare* strain Fc7 was added at three different amounts (no, low or high) to microcosms containing a lake water bacterial community. The microcosms subsequently received either a bacteriophage- or antibiotic treatment or were left untreated. Throughout the result section, treatment refers to the treatment application. Each experimental condition was replicated five times, and the bacterial community was studied over one week.

The *F. columnare* Fc7 abundance decreased in all microcosms regardless of treatment type

ASV1 was identified as the added *F. columnare* Fc7 population, and we scaled the relative ASV abundance with the bacterial density to obtain absolute ASV1 abundances. ASV1 made up, on average, 52.7±0.1% and 82.2±4.7% of the communities immediately after adding low and high amounts of *F. columnare* Fc7, respectively. Throughout the experiment, the *F. columnare* Fc7 abundance decreased in all microcosms regardless of the treatment, even in the control microcosms (**Figure 2**). However, this observation was not due to the ineffective killing of *F.*

columnare Fc7, as the decrease was more pronounced in the phage- and antibiotic-treated microcosms. For example, at day 1 in the microcosms added high amounts of *F. columnare* Fc7, we observed a 60% decline in *F. columnare* Fc7 absolute abundance in the phage- and antibiotic-treated bacterial communities compared to the control (Pairwise Wilcox test, p=0.012). Furthermore, the effectiveness of the treatments was confirmed with live-dead staining, which showed that the living population was strongly reduced when phage- or antibiotic treatment was applied (living population at in microcosms added high amounts of *F. columnare* Fc7 at day 1; Control 75%, Phage 46%, Antibiotics 61%, see Supplementary Figure 5). These declines indicate that both the phage- and antibiotic treatment effectively inactivated *F. columnare* Fc7 and that the strain was an unsuccessful invader.

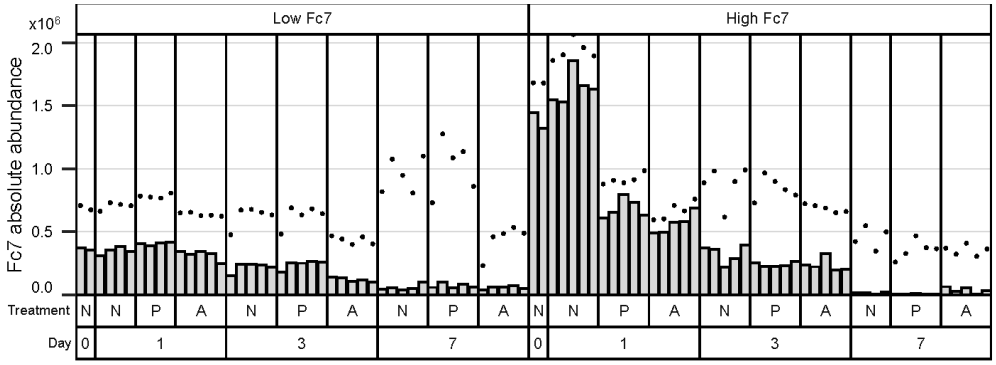


Figure 2: The absolute abundance of *Flavobacterium columnare* strain Fc7 (ASV 1) (10^6 16S rRNA gene copies/mL) in the microcosms added low and high amounts of *F. columnare* Fc7 at each sampling day. Points indicate the total bacterial density in each community (10^6 cells/mL). Abbreviations; Treatment: N = no treatment, P = Phage treatment (FCL-2), A = Antibiotic (Penicillin).

Bacterial density was impacted by the treatments

To compare how the phage- and antibiotic treatment affected the bacterial density, we fitted a mixed effect model with the bacterial density as the response variable and day, treatment (phage, antibiotic, none), amount *F. columnare* Fc7 added (no, low, high) and the interactions

252 between these as explanatory variables. Sampling day per microcosm was defined as a
253 random variable (**Figure 3**a, R^2=0.78, Supplementary Table 1). Statistical significance was
254 determined by comparing the marginal mean estimates of the phage- or antibiotic treatment
255 to the control at each day (Supplementary Table 2).

256 The effect of the phage treatment on the bacterial density varied depending on the
257 amount of *F. columnare* Fc7 added. In the microcosms without *F. columnare* Fc7 added, the
258 ratio in marginal mean bacterial density estimate between the phage treated and the control
259 increased from 1.13±0.15 (ratio±SE) at day 0 to 1.50±0.19 at day 7 ($p<0.001$, **Figure 3**b). When
260 low amounts of *F. columnare* Fc7 were added to the microcosms, the phage treatment had
261 no observable effect on the cell density, as the phage treated and the control microcosms had
262 similar cell densities over time ($p>0.05$ at all time points). However, in the microcosms added
263 high amounts of *F. columnare* Fc7, the bacterial density ratio between the phage treated and
264 control microcosms decreased from 0.78±0.10 at day 0 to 0.72±0.10 at day 7 ($p=0.03$). In fact,
265 all microcosms added high amounts of *F. columnare* Fc7 had a substantial decline in bacterial
266 density from day 5 to 7, resembling a feast-famine response. We speculate that the density
267 decline was not a result of the phage treatment but instead induced by a significant release
268 of DOM due to the death of *F. columnare* Fc7 (see Discussion).

269 The antibiotic treatment negatively impacted the bacterial density. At day 7, the
270 bacterial density was lower in the antibiotic-treated communities compared to the control,
271 with a ratio of 0.76±0.10 ($p=0.07$) in the microcosms without *F. columnare* Fc7 added,
272 0.44±0.06 ($p<0.001$) in the microcosms with low amounts of *F. columnare* Fc7 added and
273 0.76±0.10 ($p=0.09$) in the microcosms with high amounts of *F. columnare* Fc7 added (**Figure**
274 **3**b). Thus, compared to the detrimental effect of antibiotics, the impact of phage treatment
275 on bacterial density was minor.

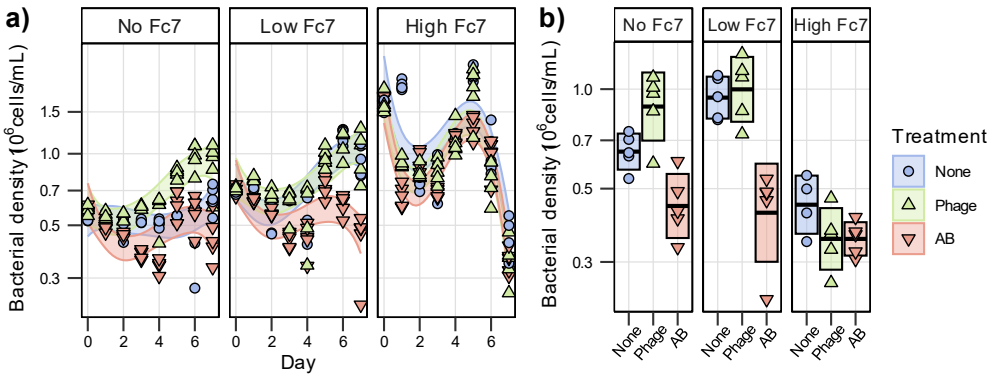


Figure 3: Changes in total bacterial density over time (y-axis is $\log_{10}$ scaled). a) The bacterial community density (10^6cells/mL) in each microcosm over time (days). The shaded area indicates the 95% confidence interval of the model prediction. b) The bacterial density at day 7. The box indicates the mean±standard deviation. Colours and shapes indicate the treatment type. Abbreviations; Treatment: None = no treatment, Phage = Phage treatment (FCL-2), AB = Antibiotic (Penicillin), No Fc7 = uninvaded, Low or High Fc7 = 24% or 190% increase in density after addition of *F. columnare* Fc7.

Phage treatment had a negligible impact on α-diversity, whereas antibiotics drastically reduced it

To evaluate the treatment effect on the α-diversity, we determined the ASV richness (**Figure 4**), Hill diversity of the first and second order, and evenness (Supplementary Figure 6). To estimate the differences between the treatments and control, we fitted a mixed effect model with ASV richness as the response variable and DPI, treatment, amount of *F. columnare* Fc7 added, and their interaction as the explanatory variables. Sampling day was included as a random effect term (R^2=0.80, Supplementary Table 3). Post-hoc comparisons are summarised in Supplementary Table 4.

We found no statistical evidence that phage treatment decreased richness. The exception was at day 3, where the richness was on average 1.12x higher in the microcosms added high amounts of *F. columnare* Fc7 receiving phage treatment than in the controls (Supplementary Table 4). The antibiotic treatment decreased richness. This reduction was

particularly evident at day 7, where the richness was on average 0.65x, 0.62x, and 0.38x lower in the antibiotic-treated than in control microcosms in the microcosms without *F. columnare* Fc7 (p-value<0.01), added low amounts of *F. columnare* Fc7 (p-value<0.001) and added high amounts of *F. columnare* Fc7 (p-value<0.001), respectively.

The observation that phage treatment had negligible effects, while antibiotics reduced the ASV richness, was also found for Hill diversity of order 1 and 2 and evenness (Supplementary Figure 6). In conclusion, the antibiotic treatment caused a loss of biodiversity. Notably, the phage treatment did not decrease α-diversity, indicating that the bacterial populations were resilient to this disturbance.

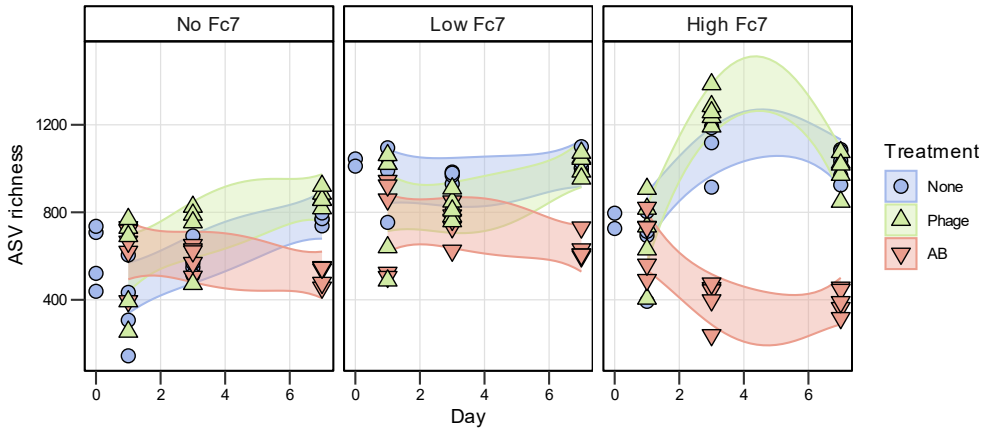


Figure 4: The ASV richness in each sample over time (days). Points are the observed richness, and the shaded area indicates the 95% confidence interval of the model prediction. Colours and shapes indicate the treatment type. Abbreviations; Treatment: None = no treatment, Phage = Phage treatment (FCL-2), AB = Antibiotic (Penicillin), No Fc7 = uninvaded, Low or High Fc7 = 24% or 190% increase in density after addition of *F. columnare* Fc7.

Bacterial community composition was similar between the phage and control treatment

The bacterial community composition was similar between the phage treatment and the control when evaluating composition at the order level (**Figure 5**) and by PCoA ordinations based on both Bray-Curtis and Sørensen similarity (**Figure 6**). After one day, there was no significant difference in the community composition between the phage treatment and untreated control, regardless of the amount of *F. columnare* Fc7 added (Bray-Curtis and Sørensen based PERMANOVA $p>0.05$, Supplementary Table 5). Thus, the data is suitable for studying the effects of the treatments on bacterial community succession.

The phage treatment did not impact the community succession in microcosms where *F. columnare* Fc7 had been added, as there were no significant differences between the phage-treated and control communities at day 7 (PERMANOVA $p>0.05$). However, for the microcosms without *F. columnare* Fc7, there was a significant difference in the community composition based on both ASV abundance (Bray-Curtis PERMANOVA $r^2=0.77$, $p=0.009$) and presence-absence (Sørensen PERMANOVA $r^2=0.26$, $p=0.009$).

Differential abundance analysis conducted on samples from day 7 identified 18 ASVs (8 genera) with a ratio of absolute abundance between the phage treatment and control below 0.2 or over 5 (Supplementary Figure 7). Of these, 14 were identified in the microcosms without *F. columnare* Fc7, six in microcosms added low amounts of *F. columnare* Fc7, and only two in the microcosms added high amounts of *F. columnare* Fc7. These differences indicate that the phage treatment impacted the uninvaded microcosms the most. Of particular interest was an ASV belonging to the genus *Aquirufa,* with an absolute abundance 227 times higher in the phage treated ($1.2x10^5 \pm 4.0x10^4$ASVs/mL) than the control (530±459ASVs/mL) microcosms without *F. columnare* Fc7 added. Despite significant differences, the average

Bray-Curtis similarity between the phage-treated and control communities only changed slightly from 0.77±0.14 at day 1 to 0.67±0.12 at day 7 (**Figure 6**a). Thus, we conclude that the phage treatment had no impact on community composition and succession when the phages bacterial host was added, but that minor changes were induced when the phage host was absent.

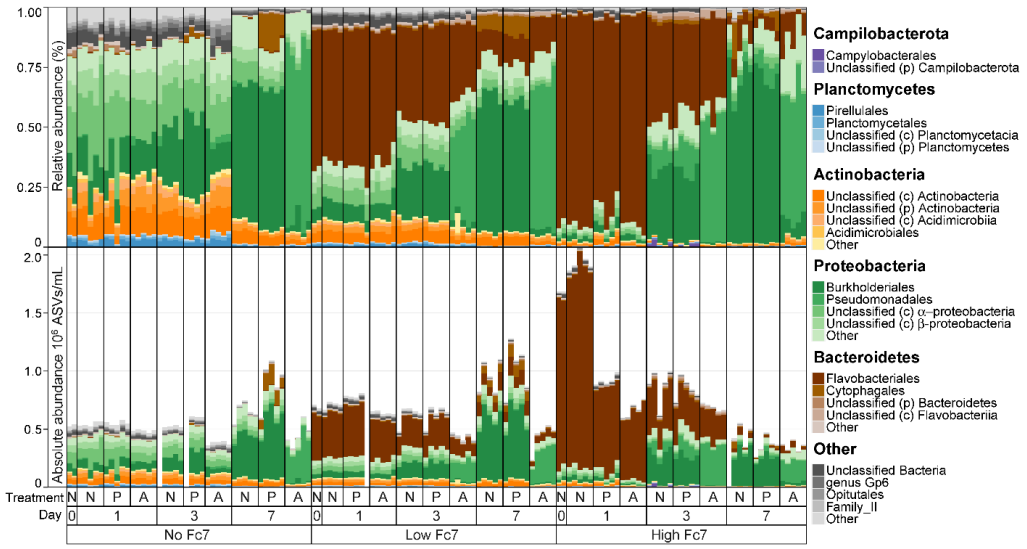


Figure 5: The bacterial community composition at the order level grouped according to bacterial phyla. Only the four most abundant orders are shown, and the rest are grouped as Other. Community composition was evaluated as the relative- (upper panel, %) and absolute abundance (bottom panel, 10^6 ASV copies/mL). Abbreviations; Treatment: N = no treatment, P = Phage treatment (FCL-2), A = Antibiotic (Penicillin), No Fc7 = uninvaded, Low or High Fc7 = 24% or 190% increase in density after addition of *F. columnare* Fc7.

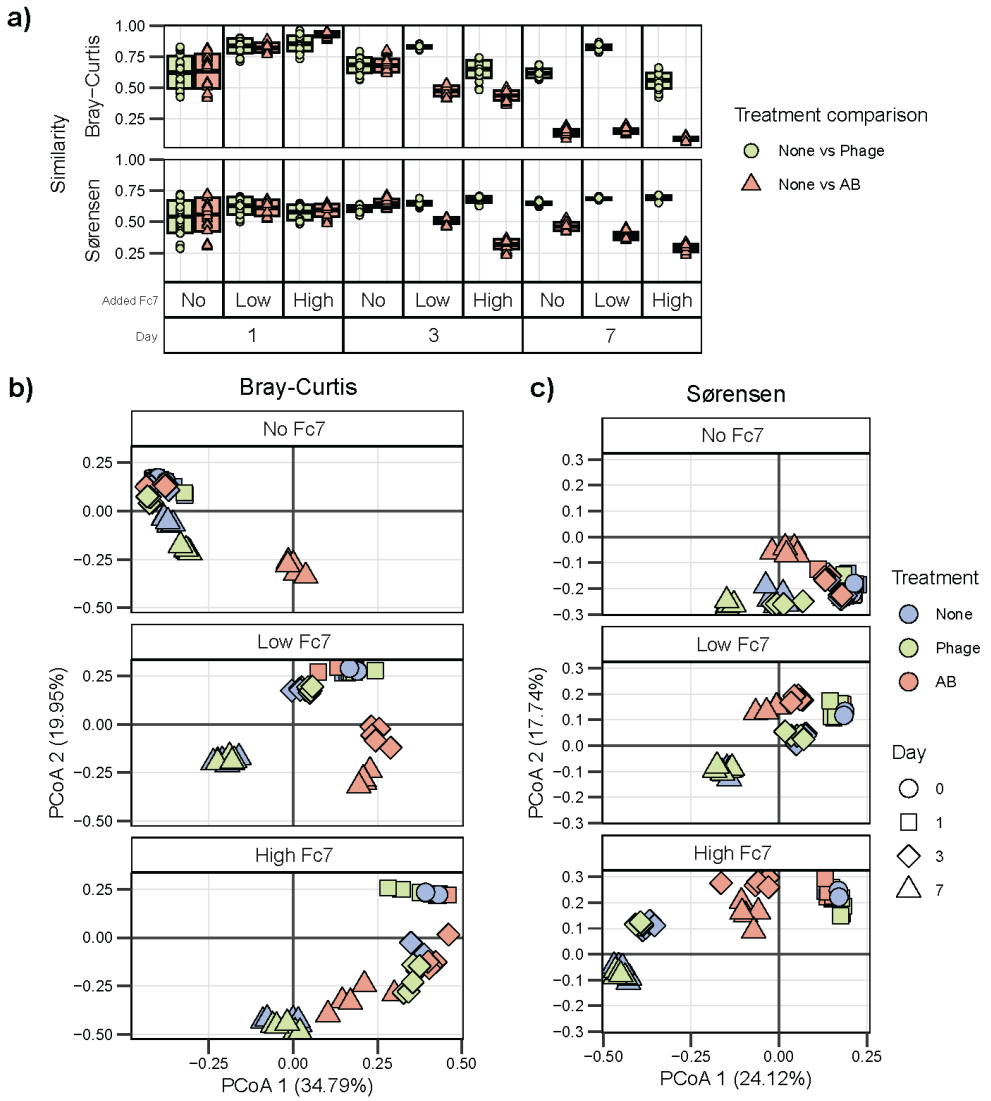


Figure 6: Bray-Curtis and Sørensen similarity between samples indicated that communities changed over time. A) Bray-Curtis (upper) and Sørensen similarity (lower panel) comparing the control (none) and the treatments phage or antibiotic, at each sampling day and level of *F. columnare Fc7* added. Colour and shape indicate comparison. B and c) Bray-Curtis and Sørensen based PcoA ordinations of samples taken at day 0, 1, 3 and 7. Each plot is a single ordination but is separated based on the amount of *F. columnare* Fc7 added for clarity. Colours and shapes indicate the treatment type and sampling day. Abbreviations; Treatment: None = no treatment, Phage = Phage treatment (FCL-2), AB = Antibiotic (Penicillin), No Fc7 = uninvaded, Low or High Fc7 = 24% or 190% increase in density after addition of *F. columnare* Fc7.

Antibiotic treatment caused a significant disturbance event in the community

In contrast to the phage treatment, antibiotics caused the community composition to change significantly compared to the control microcosms (**Figure 6**). At day 7, the community composition significantly differed between the antibiotic-treated and control microcosms (PERMANOVA, r^2 range 0.52-0.84, $p<0.05$ for both Bray-Curtis and Sørensen). These changes are evident in the average similarity between communities from the antibiotic-treated and control microcosms. There was a 6x reduction in Bray-Curtis similarity (0.79±0.15 at day 1, 0.13±0.035 at day 7) and a 1.5x reduction in Sørensen similarity (0.59±0.095 at day 1, 0.38±0.078 at day 7) (**Figure 6**a).

Differential abundance analysis identified 122 ASVs (29 genera) with a ratio in absolute abundance between the antibiotic-treated and control communities below 0.2 or over 5 (Supplementary Figure 7). Thus, there were 4.2x more ASVs with such a substantial difference in the antibiotic-treated than the phage-treated microcosms. Interestingly, of the 122 ASVs, all ASVs classified as *β-proteobacteria* (68 ASVs) had higher absolute abundances in the control microcosms, while all classified as *γ-proteobacteria* (11 ASVs) had higher absolute abundances in the antibiotic-treated microcosms. These 11 ASVs belonged to the genus *Pseudomonas* which contains many pathogenic bacterial strains [49].

In conclusion, the antibiotic treatment caused significant disturbances that the bacterial communities did not recover from after seven days and caused a bloom of *Pseudomonas*.

Phage treatment did not affect the bacterial community assembly, while antibiotics caused a shift from selection to drift

We investigated the community assembly within each experimental condition by calculating the change in similarity between replicate microcosms over time (i.e. similarity rate) as described in Gundersen et al. 2021 [40]. Increasing similarity rates indicate that the deterministic process selection dominates community assembly. In contrast, decreasing similarity rates indicate an increased contribution of the stochastic process ecological drift.

We used this assembly framework with both the Bray-Curtis and Sørensen similarity (**Figure 7**, Supplementary Figure 8). We first examined the microcosms without *F. columnare* Fc7 added to evaluate the effect of phage- and antibiotic treatment on community assembly when no phage host was present (i.e. not considering the effect of adding *F. columnare* Fc7). The phage treated and control microcosms had comparable positive similarity rates, with the average varying between 0.033-0.040/day for the Bray-Curtis and 0.024-0.025/day for the Sørensen similarity rate. These positive rates indicated that the communities in replicate microcosms became more similar over time and were thus primarily structured by selection. On the other hand, the antibiotic-treated microcosms had a negative similarity rate based on Bray-Curtis (−0.049/day), while the Sørensen-based was slightly positive (0.003/day). Thus, when antibiotics were added, the community composition was structured by drift, with some selection at the ASV inventory level (**Figure 7**).

Next, we determined if treatment and amount of *F. columnare* Fc7 added combined affected community assembly. At each level of added *F. columnare* Fc7, the Bray-Curtis and Sørensen similarity rates of the phage-treated microcosms were not significantly different from the control microcosms. However, the average similarity rate decreased with increasing amounts of *F. columnare* Fc7 added for both the phage-treated and the control microcosms.

This effect of *F. columnare* Fc7 amounts added was not observed for the antibiotic-treated microcosms, where the Bray-Curtis similarity rate was clearly negative regardless of amounts of *F. columnare* Fc7 added (no = -0.049, low = -0.023, high = -0.070). Thus, the treatment (phage and antibiotic) and amounts of *F. columnare* Fc7 did not have an additive effect.

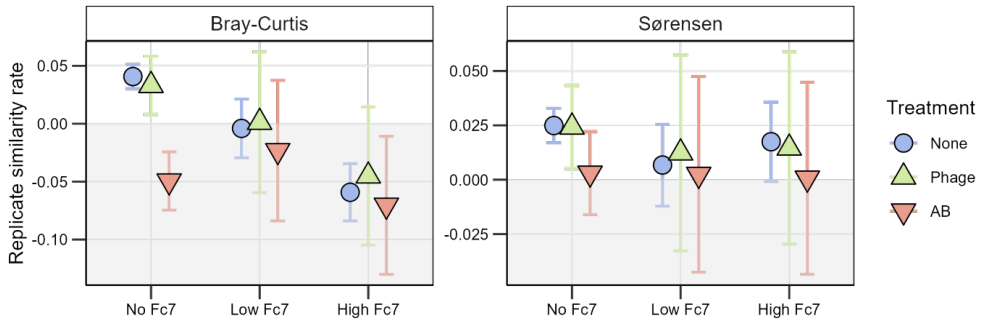


Figure 7: Estimated replicate similarity rate for experimental group based on a) Bray-Curtis and b) Sørensen similarity. The error bars are the 95% confidence interval of the estimated similarity rate of change. Colours and shapes indicate the treatment type. Abbreviations; Treatment: None = no treatment, Phage = Phage treatment (FCL-2), AB = Antibiotic (Penicillin), No Fc7 = uninvaded, Low or High Fc7 = 24% or 190% increase in density after addition of *F. columnare* Fc7.

Discussion

Comprehending the ecological impacts of phage therapy is vital due to the increased interest in applying this technology [18, 50]. A concern is a lack of understanding of how phages impact the resident bacterial community. From one perspective, it is desirable to use therapeutic agents that minimally impact the microbiome when treating bacterial infections in animals or humans to ensure stability in the bacterial community. Currently, this knowledge gap hinders clinical approvals for using phages in humans due to the potential disturbance it can cause to the microbiome [51]. Furthermore, phages are introduced into ecosystems through, for example, water released from aquaculture facilities [18] and when sprayed over agricultural fields [50]. Thus, we must elucidate how the bacterial communities in such phage-receiving ecosystems respond to phage exposure. As such, the study of the impacts of phage therapy is not just a scientific pursuit but also relevant for society with the aim to ensure safety and sustainability of ecosystems.

Most studies investigating the effect of phage treatment on the properties of the resident community have been performed *in situ* (e.g. human and animal gut). Although more realistic, these ecosystems contain many unknown or uncontrollable variables that can mask changes in the resident bacterial community [52]. Our experiment aimed to reduce the environmental complexity and minimise within-group variability by bringing a planktonic lake community into controlled laboratory conditions. The observed community composition was highly similar between biological replicates at day 1, indicating that our goal of creating replicate communities was successful.

The aim of this study was to investigate how phage treatment affected community structure. We were particularly interested in understanding how the amount of the bacterial

host of the phage affected the outcome. The release of resources, such as mineral nutrients and DOM, is proportional to the number of hosts lysed. We hypothesized that over a threshold, lysis products would induce a change in the structure of the resident bacterial community. We therefore introduced the phage host *F. columnare* Fc7 at two different levels.

As the added *F. columnare* Fc7 did not establish in the system, the invasion was unsuccessful. Thus, also the resident community in the control and treatment groups without addition of phage experienced an increase in available resources due to the death of *F. columnare* Fc7. To account for this unsuccessful invasion our analytical approach was to compare the treated and control microcosms with the same level of *F. columnare* Fc7 added. These comparisons reflect the effects of adding phage or antibiotic to increase the death rate of a declining population. Had the *F. columnare* Fc7 population successfully established in the control microcosms, the results may have been different.

Furthermore, in the current experiment, the *F. columnare* Fc7 invasion was performed only an hour before the treatments were applied. It is unrealistic that *F. columnare* Fc7 formed any meaningful interactions with the resident community in that timeframe. Consequently, we could not evaluate the possible changes phage treatment induces in the bacterial interaction networks. It has been demonstrated that phages can result in cascading effects in the interaction network of a 10-species synthetic community [53]. Therefore, future studies should investigate the impact of removing an established population from the resident community.

Our comprehensive study of the bacterial communities showed that phage treatment had a marginal biological effect on community properties such as density, α-diversity, composition, succession, and assembly, compared to the control. Most analyses showed no statistically significant changes between the phage-treated and control microcosms. When *F.*

columnare Fc7 was added to the community, statistically significant effects of phage treatment were detected in only a few cases, indicating that the treatment had little effect on community structure. We observed that the bacterial density in the phage-treated microcosms was halved after 1 day compared to the control, indicating that the phage was successful in targeting and lysing the *F. columnare* Fc7 cells. Quantification of the free phage population would have provided a better understanding of the phage dynamics in this experimental setup. On day 3, ASV richness was 12% higher in the phage-treated microcosms compared to the control microcosms where high levels of *F. columnare* Fc7 were added. However, this was the only occasion when there was a significant difference in diversity. Therefore, when evaluating the totality of the analyses performed, it appears that phage treatment induces only minor to negligible effects on community characteristics when the phage host is present.

Furthermore, we found no evidence for a greater effect of phage treatment on community structure, when more hosts were added. In general, no significant differences were detected between the phage-treated and control microcosms. This applies for both low and high levels of *F. columnare* Fc7 added. This is probably related to the fact that *F. columnare* Fc7 also died in the control microcosms, and future studies should challenge these findings with a more successful invader.

In contrast to our hypothesis, we observed the largest impact of the phage treatment on the bacterial community characteristics when no phage host (i.e. *F. columnare* Fc7) was present. When the microcosms without *F. columnare* Fc7 were added phages the density was on average 38.4% higher than in the control, and we observed a significant change in the community succession that resulted in changes in ASV inventory and relative abundance compared to the control. The changes were mainly driven by a substantial increase in a single

ASV classified as *Aquirufa*. This genus is part of the order *Cytophagales,* known to be efficient degraders of biopolymers such as proteins, DNA and RNA [54]. It is known that phages can function as a substrate for heterotrophic bacterial growth [34]. Noble et al. 1999 observed that bacterial density increased after adding a phage cocktail to a bacterial community. The authors concluded that the phage particles stimulated the growth of non-infected heterotrophic bacteria [55]. In a follow-up study, they radiolabelled viral components and subsequently found them incorporated into the bacterial biomass [56]. Phages are essential in the microbial loop by increasing DOM turnover through the lysis of bacteria (i.e. viral shunt). Our observations indicate that some non-target resident bacteria benefit from viral decay. Thus, exploring how viral decay contributes to the microbial loop would be fascinating and appears to be a knowledge gap.

We observed that the bacterial density fluctuated in a feast-famine response manner in the microcosms added high amounts of *F. columnare* Fc7, regardless of the treatment. The feast-famine response occurs when communities experience a surge in available resources, leading to an increase in density, followed by a decline due to famine after resource consumption [57]. The *F. columnare* Fc7 population declined drastically during the first days. When bacteria lyse, DOM is released, which can be consumed by the resident community [58, 59]. From day 3 to 5, we observed a doubling in cell density, which might be explained by a feast on DOM released from lysed *F. columnare* Fc7 cells. Following the depletion of DOM, we observed a substantial 4.5-fold decrease in density. This decrease likely occurred as the carrying capacity of the system could not sustain the peak in population density, which led to famine-induced mortality. We speculate that this famine response was stronger in the phage treatment, possibly due to the initial pulse in resources at day 1 due to lysis of *F. columnare* Fc7. This stronger response may explain why we observed a 21% lower bacterial density on

512 day 7 in the microcosms added high amounts of *F. columnare* Fc7 and phage treatment as
513 compared to the control. Our understanding of such feast-famine responses in controlled
514 experimental settings is still poor and further investigation of these dynamics would be
515 beneficial to the field.

516 Increased amounts of *F. columnare* Fc7 added was accompanied by a shift from
517 selection to ecological drift dominating the community assembly. Zhou et al. 2014
518 hypothesised that nutrient disturbances should enhance stochastic community assembly due
519 to reduced niche selection and growth of the rare biosphere [60]. They found evidence for
520 their hypothesis using vegetable oil as a nutrient spike. Through 16S rRNA gene sequencing
521 and flow cytometry, we showed that the relative and absolute abundance of *F. columnare*
522 Fc7 declined, possibly leading to a substantial increase in DOM. Thus, we support their
523 hypothesis by showing that dead bacterial cells increase the contribution of stochastic
524 processes. Increased stochasticity results in more unpredictable changes at the community
525 level and, consequently, replicate microcosms diverge from each other. Hence, researchers
526 should carefully consider how much of the invader they will add when planning invasion
527 studies. Our observations indicate that too high invader concentrations can result in a feast-
528 famine response and increased ecological drift.

529 The antibiotic treatment functioned well as a positive disturbance control. The
530 antibiotic treatment caused a substantial decrease in bacterial density, reduced α-diversity,
531 significantly changed community composition and enhanced stochastic community assembly.
532 These characteristics are indicative that the antibiotic treatment caused a severe disturbance.
533 We conclude that the resident community had little resistance to this disturbance, as it was
534 strong enough to push the community out of its stable state [61, 62].

The bacterial density declined significantly in the antibiotic-treated communities compared to the control. This decline and a loss in ASV richness indicate that several bacterial populations died. This bacterial death has two effects: increased DOM release and reduced niche competition. Both effects are characteristic for conditions that select for opportunistic r-strategic bacteria that respond to high resource availability by rapidly increasing their growth rate leading to a numeric response in cell density [59]. Most pathogenic bacteria are classified as r-strategic bacteria [59]. Intriguingly, the antibiotic treatment significantly increased ASVs classified as *Pseudomonas* (*γ-proteobacteria*). This genus is associated with r-strategic organisms [63] and contains many pathogenic bacteria, such as *P. aeruginosa* and *P. fluorescens* [49]. Thus, our observations show that the antibiotic treatment created an environment that allowed opportunistic bacteria to bloom. This discovery is concerning due to the possibility that antibiotic treatment may result in dysbiosis and selection for antibiotic-resistant pathogenic bacteria.

Conclusions

Our study investigated the impact of phage treatment on resident community structure. This study is the first to explore how the density of the phage host impacts the effects of phage treatment on community structure. The amount of host (*F. columnare* Fc7) had an impact. We found that a single FCL-2 phage treatment had a negligible impact on community diversity and composition when the host *F. columnare* Fc7 was added to the community. Interestingly, we observed significant effects of phage addition when *F. columnare* Fc7 was absent, mostly driven by an increase in the abundance of *Aquirufa sp*. Further investigations should explore the underlying mechanisms for this observation. Nevertheless, when changes due to the phage treatment were observed, they were minor

compared to the detrimental impacts of the antibiotic treatment. Antibiotics resulted in a substantial decline in bacterial density and α-diversity, altered the community composition and triggered a bloom of opportunistic bacteria. Such drastic changes were not observed for phage treatment in the presence of the bacterial host. These findings are relevant for treatment of humans and for industries such as aquaculture, agriculture, and wastewater management, as they benefit from stable and functional microbial communities. As the phage treatment induced only minor changes to community structure, our observations indicate that phage therapy is a safer and superior alternative to antibiotics for therapeutic use in host microbiomes, such as the human gut.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and material

All data generated from flow cytometry (FSC files) are available at https://figshare.com/projects/Flow_cytometry_data/152886. The Illumina sequencing reads are deposited at the European Nucleotide Archive (accession number PRJEB59722).

Competing interests

The authors declare that they have no competing interests.

Funding

The Faculty of Natural Science at NTNU funded this research as a PhD scholarship to MSG and AWF.

Authors' contributions

MSG and AWF conceptualised the research idea. MSG, AWF, OV and IB designed and planned the research. MSG and AWF conducted the experiment and processed samples. MSG analysed the data and wrote the first draft of the manuscript. MSG, AWF, OV, and IB contributed to developing the hypothesis and revising the manuscript.

Acknowledgements

We kindly thank David Perez-Pascual and Jean-Marc Ghigo (Institute Pasteur, France) for providing *Flavobacterium columnare* strain Fc7 and Lotta-Riina Sundberg (University of Jyväskylä, Finland) for providing the phage isolate FCL-2. We also appreciate Toan Vo and Jenny Poppe (Norwegian University of Science and Technology, Norway) for helping us enrich the phage stock. Further, we thank Andrew Morris (University of Oregon, USA) for his input and discussions on the statistical analysis of the dataset. We would also like to thank two anonymous reviewers for helping improve this manuscript.

Authors' information

Madeleine S. Gundersen is a PhD candidate studying how disturbances affect bacterial community characteristics and assembly.

Alexander W. Fiedler is a PhD candidate researching how the microbiome and bacterial infections impact the bacterial community of *Atlantic salmon* and the microbiome-host interplay.

Ingrid Bakke is a professor in environmental biotechnology at NTNU. Her research focuses on fish-microbiome interactions, microbial management in land-based aquaculture and biological water treatment.

Olav Vadstein is a professor in microbial ecology at NTNU. His focus is on the structure and function of natural and built aquatic ecosystems, host-microbiome interactions and microbial management in fish, and microbial water quality in aquaculture.

All authors are employed at the Department of Biotechnology and Food Science at the Norwegian University of Science and Technology (NTNU).

References

1. Spellberg B. The future of antibiotics. Crit Care 2014; https://doi.org/10.1186/cc13948.
2. Llor C, Bjerrum L. Antimicrobial resistance: risk associated with antibiotic overuse and initiatives to reduce the problem. Ther Adv Drug Saf 2014; https://doi.org/10.1177/2042098614554919.
3. Grenni P, Ancona V, Barra Caracciolo A. Ecological effects of antibiotics on natural ecosystems: A review. Microchemical Journal 2018; https://doi.org/10.1016/j.microc.2017.02.006.
4. Chanishvili N. Phage therapy - history from Twort and d'Herelle through Soviet experience to current approaches. Adv Virus Res 2012; https://doi.org/10.1016/B978-0-12-394438-2.00001-3.
5. Górski A, Międzybrodzki R, Węgrzyn G, Jończyk-Matysiak E, Borysowski J, Weber-Dąbrowska B. Phage therapy: Current status and perspectives. Med Res Rev 2020; https://doi.org/10.1002/med.21593.
6. Salmond GPC, Fineran PC. A century of the phage: past, present and future. Nature Reviews Microbiology 2015 13:12 2015; https://doi.org/10.1038/nrmicro3564.
7. Yassour M, Vatanen T, Siljander H, Hämäläinen AM, Härkönen T, Ryhänen SJ, et al. Natural history of the infant gut microbiome and impact of antibiotic treatment on bacterial strain diversity and stability. Sci Transl Med 2016; https://doi.org/10.1126/scitranslmed.aad0917.
8. Dubourg G, Lagier JC, Robert C, Armougom F, Hugon P, Metidji S, et al. Culturomics and pyrosequencing evidence of the reduction in gut microbiota diversity in patients with broad-spectrum antibiotics. Int J Antimicrob Agents 2014; https://doi.org/10.1016/j.ijantimicag.2014.04.020.
9. Gao Q, Gao S, Bates C, Zeng Y, Lei J, Su H, et al. The microbial network property as a bio-indicator of antibiotic transmission in the environment. Science of The Total Environment 2021; https://doi.org/10.1016/j.scitotenv.2020.143712.
10. Blaser M. Stop the killing of beneficial bacteria. Nature 2011 476:7361 2011; https://doi.org/10.1038/476393a.
11. Maier L, Goemans C V, Wirbel J, Kuhn M, Eberl C, Pruteanu M, et al. Unravelling the collateral damage of antibiotics on gut bacteria. Nature 2022; https://doi.org/10.1038/s41586-021-03986-2.
12. Patangia D V., Anthony Ryan C, Dempsey E, Paul Ross R, Stanton C. Impact of antibiotics on the human microbiome and consequences for host health. Microbiologyopen 2022; https://doi.org/10.1002/mbo3.1260.
13. Rios AC, Moutinho CG, Pinto FC, Del Fiol FS, Jozala A, Chaud M V., et al. Alternatives to overcoming bacterial resistances: State-of-the-art. Microbiol Res 2016; https://doi.org/10.1016/j.micres.2016.04.008.
14. Bergh Ø, Børsheim KY, Bratbak G, Heldal M. High abundance of viruses found in aquatic environments. Nature 1989 340:6233 1989; https://doi.org/10.1038/340467a0.
15. Gill JJ, Hyman P. Phage Choice, Isolation, and Preparation for Phage Therapy. Curr Pharm Biotechnol 2010; https://dx.doi.org/10.2174/138920110790725311.
16. Olsen Y, Andersen T, Gismervik I, Vadstein O. Protozoan and metazoan zooplankton-mediated carbon flows in nutrient-enriched coastal planktonic communities. Mar Ecol Prog Ser 2007; 331: 67–83.

17. Kowalska JD, Kazimierczak J, Sowińska PM, Wójcik EA, Siwicki AK, Dastych J. Growing Trend of Fighting Infections in Aquaculture Environment—Opportunities and Challenges of Phage Therapy. Antibiotics 2020; https://doi.org/10.3390/antibiotics9060301.
18. Culot A, Grosset N, Gautier M. Overcoming the challenges of phage therapy for industrial aquaculture: A review. Aquaculture 2019; https://doi.org/10.1016/j.aquaculture.2019.734423.
19. Madhusudana Rao B, Lalitha K V. Bacteriophages for aquaculture: Are they beneficial or inimical. Aquaculture 2015; https://doi.org/10.1016/j.aquaculture.2014.11.039.
20. Sinha A, Li Y, Mirzaei MK, Shamash M, Samadfam R, King IL, et al. Transplantation of bacteriophages from ulcerative colitis patients shifts the gut bacteriome and exacerbates the severity of DSS colitis. Microbiome 2022 10:1 2022; https://doi.org/10.1186/s40168-022-01275-2.
21. Juan CAPD, Gómez-Verduzco G, Márquez-Mota CC, Téllez-Isaías G, Kwon YM, Cortés-Cuevas A, et al. Productive Performance and Cecum Microbiota Analysis of Broiler Chickens Supplemented with β-Mannanases and Bacteriophages—A Pilot Study. Animals 2022; https://doi.org/10.3390/ani12020169.
22. Zhao H, Li Y, Lv P, Huang J, Tai R, Jin X, et al. Salmonella Phages Affect the Intestinal Barrier in Chicks by Altering the Composition of Early Intestinal Flora: Association With Time of Phage Use. Front Microbiol 2022; https://doi.org/10.3389/fmicb.2022.947640.
23. Ahasan MS, Kinobe R, Elliott L, Owens L, Scott J, Picard J, et al. Bacteriophage versus antibiotic therapy on gut bacterial communities of juvenile green turtle, Chelonia mydas. Environ Microbiol 2019; https://doi.org/10.1111/1462-2920.14644.
24. Lorenzo-Rebenaque L, Casto-Rebollo C, Diretto G, Frusciante S, Rodríguez JC, Ventero MP, et al. Examining the effects of Salmonella phage on the caecal microbiota and metabolome features in Salmonella-free broilers. Front Genet 2022; https://doi.org/10.3389/fgene.2022.1060713.
25. Hong Y, Thimmapuram J, Zhang J, Collings CK, Bhide K, Schmidt K, et al. The impact of orally administered phages on host immune response and surrounding microbial communities. Bacteriophage 2016; https://doi.org/10.1080/21597081.2016.1211066.
26. Grubb DS, Wrigley SD, Freedman KE, Wei Y, Vazquez AR, Trotter RE, et al. PHAGE-2 Study: Supplemental Bacteriophages Extend Bifidobacterium animalis subsp. lactis BL04 Benefits on Gut Health and Microbiota in Healthy Adults. Nutrients 2020, Vol 12, Page 2474 2020; https://doi.org/10.3390/nu12082474.
27. McCallin S, Alam Sarker S, Barretto C, Sultana S, Berger B, Huq S, et al. Safety analysis of a Russian phage cocktail: From MetaGenomic analysis to oral application in healthy human subjects. Virology 2013; https://doi.org/10.1016/j.virol.2013.05.022.
28. Sarker SA, Berger B, Deng Y, Kieser S, Foata F, Moine D, et al. Oral application of Escherichia coli bacteriophage: safety tests in healthy and diarrheal children from Bangladesh. Environ Microbiol 2017; https://doi.org/10.1111/1462-2920.13574.
29. Zhao J, Liu Y, Xiao C, He S, Yao H, Bao G. Efficacy of phage therapy in controlling rabbit colibacillosis and Changes in cecal microbiota. Front Microbiol 2017; https://doi.org/10.3389/fmicb.2017.00957.
30. Galtier M, De Sordi L, Maura D, Arachchi H, Volant S, Dillies MA, et al. Bacteriophages to reduce gut carriage of antibiotic resistant uropathogens with low impact on microbiota composition. Environ Microbiol 2016; https://doi.org/10.1111/1462-2920.13284.

31. Hu YOO, Hugerth LW, Bengtsson C, Alisjahbana A, Seifert M, Kamal A, et al. Bacteriophages Synergize with the Gut Microbial Community To Combat Salmonella. mSystems 2018; https://doi.org/10.1128/mSystems.00119-18.
32. Pereira C, Silva YJ, Santos AL, Cunha Â, Gomes NCM, Almeida A. Bacteriophages with potential for inactivation of fish pathogenic bacteria: survival, host specificity and effect on bacterial community structure. Mar Drugs 2011; https://doi.org/10.3390/md9112236.
33. Braga LPP, Spor A, Kot W, Breuil MC, Hansen LH, Setubal JC, et al. Impact of phages on soil bacterial communities and nitrogen availability under different assembly scenarios. Microbiome 2020; https://doi.org/10.1186/s40168-020-00822-z.
34. Kirchman DL. Processes in microbial ecology. 2012. Oxford University Press, Oxford.
35. Perez-Pascual D, Vendrell-Fernandez S, Audrain B, Bernal-Bayard J, Patiño-Navarrete R, Petit V, et al. Gnotobiotic rainbow trout (Oncorhynchus mykiss) model reveals endogenous bacteria that protect against Flavobacterium columnare infection. PLoS Pathog 2021; https://doi.org/10.1371/journal.ppat.1009302.
36. Laanto E, Sundberg LR, Bamford JKH. Phage specificity of the freshwater fish pathogen Flavobacterium columnare. Appl Environ Microbiol 2011; https://doi.org/10.1128/AEM.05574-11.
37. Glonti T, Pirnay J-P. In Vitro Techniques and Measurements of Phage Characteristics That Are Important for Phage Therapy Success. Viruses 2022, Vol 14, Page 1490 2022; https://doi.org/10.3390/v14071490.
38. Vestrum RI, Attramadal KJK, Vadstein O, Gundersen MS, Bakke I. Bacterial community assembly in Atlantic cod larvae (Gadus morhua): contributions of ecological processes and metacommunity structure. FEMS Microbiol Ecol 2020; https://doi.org/10.1093/femsec/fiaa163.
39. Edgar RC. Search and clustering orders of magnitude faster than BLAST. Bioinformatics 2010; https://doi.org/10.1093/bioinformatics/btq461.
40. Gundersen MS, Morelan IA, Andersen T, Bakke I, Vadstein O. The effect of periodic disturbances and carrying capacity on the significance of selection and drift in complex bacterial communities. ISME Communications 2021; https://doi.org/10.1038/s43705-021-00058-4.
41. R Core Team. R: A language and environment for statistical computing. 2022. R Foundation for Statistical Computing, Vienna, Austria.
42. Hill MO. Diversity and Evenness: A Unifying Notation and Its Consequences. Ecology 1973; https://doi.org/10.2307/1934352.
43. Bray JR, Curtis JT. An Ordination of the Upland Forest Communities of Southern Wisconsin. Ecol Monogr 1957; https://doi.org/10.2307/1942268.
44. Oksanen J, Simpson G, Blanchet G, Kindt R, Legendre P, Minchin P, et al. vegan: Community Ecology Package. 2022.
45. Martin BD, Witten D, Willis AD. Modeling microbial abundances and dysbiosis with beta-binomial regression. Annals of Applied Statistics 2020; https://doi.org/10.1214/19-AOAS1283.
46. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 2014; https://doi.org/10.1186/s13059-014-0550-8.
47. Lin H, Peddada S Das. Analysis of compositions of microbiomes with bias correction. Nature Communications 2020 11:1 2020; https://doi.org/10.1038/s41467-020-17041-7.

48. Nearing JT, Douglas GM, Hayes MG, MacDonald J, Desai DK, Allward N, et al. Microbiome differential abundance methods produce different results across 38 datasets. Nature Communications 2022 13:1 2022; https://doi.org/10.1038/s41467-022-28034-z.
49. Peix A, Ramírez-Bahena MH, Velázquez E. Historical evolution and current status of the taxonomy of genus Pseudomonas. Infection, Genetics and Evolution 2009; https://doi.org/10.1016/j.meegid.2009.08.001.
50. Połaska M, Sokołowska B. Bacteriophages—a new hope or a huge problem in the food industry. AIMS Microbiol 2019; https://doi.org/10.3934/microbiol.2019.4.324.
51. Ganeshan SD, Hosseinidoust Z. Phage Therapy with a Focus on the Human Microbiota. Antibiotics 2019; https://doi.org/10.3390/antibiotics8030131.
52. Konopka A. What is microbial community ecology? The ISME Journal 2009 3:11 2009; https://doi.org/10.1038/ismej.2009.88.
53. Hsu BB, Gibson TE, Yeliseyev V, Liu Q, Lyon L, Bry L, et al. Dynamic Modulation of the Gut Microbiota and Metabolome by Bacteriophages in a Mouse Model. Cell Host Microbe 2019; https://doi.org/10.1016/j.chom.2019.05.001.
54. Reichenbach H. The Order Cytophagales. The Prokaryotes 2006; https://doi.org/10.1007/0-387-30747-8_20.
55. Noble RT, Middelboe M, Fuhrman JA. Effects of viral enrichment on the mortality and growth of heterotrophic bacterioplankton. Aquatic Microbial Ecology 1999; 18: 1–13.
56. Noble RT, Fuhrman JA. Breakdown and microbial uptake of marine viruses and other lysis products. Aquatic Microbial Ecology 1999; 20: 1–11.
57. Himeoka Y, Mitarai N. Dynamics of bacterial populations under the feast-famine cycles. Phys Rev Res 2020; https://doi.org/10.1103/PhysRevResearch.2.013372.
58. Hess-Erga OK, Blomvågnes-Bakke B, Vadstein O. Recolonization by heterotrophic bacteria after UV irradiation or ozonation of seawater; a simulation of ballast water treatment. Water Res 2010; https://doi.org/10.1016/j.watres.2010.06.059.
59. Vadstein O, Attramadal KJK, Bakke I, Olsen Y. K-Selection as Microbial Community Management Strategy: A Method for Improved Viability of Larvae in Aquaculture. Front Microbiol 2018; https://doi.org/10.3389/fmicb.2018.02730.
60. Zhou J, Deng Y, Zhang P, Xue K, Liang Y, Nostrand JD Van, et al. Stochasticity, succession, and environmental perturbations in a fluidic ecosystem. PNAS 2014; https://doi.org/10.1073/pnas.1324044111.
61. Shade A, Peter H, Allison SD, Baho DL, Berga M, Bürgmann H, et al. Fundamentals of Microbial Community Resistance and Resilience. Front Microbiol 2012; https://doi.org/10.3389/fmicb.2012.00417.
62. Scheffer M, Carpenter S, Foley JA, Folke C, Walker B. Catastrophic shifts in ecosystems. Nature 2001 413:6856 2001; https://doi.org/10.1038/35098000.
63. Ho A, Di Lonardo DP, Bodelier PLE. Revisiting life strategy concepts in environmental microbial ecology. FEMS Microbiol Ecol 2017; https://doi.org/10.1093/femsec/fix006.

Paper VI

Phage therapy minimally affects the water microbiota in an Atlantic salmon (*Salmo salar*) rearing system while still preventing infection

scientific reports

Check for updates

OPEN

Phage therapy minimally affects the water microbiota in an Atlantic salmon (*Salmo salar*) rearing system while still preventing infection

Alexander W. Fiedler, Madeleine S. Gundersen, Toan P. Vo, Eivind Almaas, Olav Vadstein & Ingrid Bakke✉

Excessive usage of antibiotics threatens the bacterial diversity in the microbiota of animals. An alternative to antibiotics that has been suggested to not disturb the microbiota is (bacterio)phage therapy. In this study, we challenged germ-free and microbially colonized yolk sac fry of Atlantic salmon with *Flavobacterium columnare* and observed that the mere presence of a microbiota protected the fish against lethal infection. We then investigated the effect of phage- or oxytetracycline treatment on fish survival and rearing water bacterial community characteristics using 16S rRNA gene amplicon sequencing. Phage treatment led to an increased survival of *F. columnare*-challenged fish and reduced the relative amounts of the pathogen in the water microbiota. In the absence of *F. columnare*, phage treatment did not affect the composition or the α-diversity of the rearing water microbiota. In the presence of the phage's host, phage treatment induced minor changes to the bacterial community composition, without affecting the α-diversity. Surprisingly, oxytetracycline treatment had no observable effect on the water microbiota and did not reduce the relative abundance of *F. columnare* in the water. In conclusion, we showed that phage treatment prevents mortality while not negatively affecting the rearing water microbiota, thus suggesting that phage treatment may be a suitable alternative to antibiotics. We also demonstrated a protective effect of the microbiota in Atlantic salmon yolk sac fry.

Animal hosts benefit greatly from microbial colonization which provides nutrients for growth, supports normal development[1,2] and are assumed to protect against infections[3–7]. However, experimental proof of this protective function is scarce, especially outside of model organisms. Furthermore, these beneficial functions are threatened by a decline in microbial diversity, leaving the organism more vulnerable to infection[8,9]. One major reason for this decline is the use of antibiotics[10]. Antibiotics disturb the microbiota, leading to decreased protection against new infections or to immediate secondary infections[11]. Furthermore, the overuse of antibiotics has resulted in the selection of antibiotic-resistant strains, thereby diminishing the efficacy of these drugs[12].

One of the most used antibiotics for animal production in Europe is oxytetracycline (OTC)[13]. This broad-spectrum antibiotic binds to the 30S ribosomal subunit which inhibits bacterial protein synthesis[14,15]. Due to its widespread industrial use in aquaculture, OTC has been found in ecosystems close to fish farms, such as in sediments, water bodies and aquatic organisms[16–18]. Already low concentrations of OTC can negatively affect aquatic organisms and their microbiota[19–22] and can lead to an increase in antibiotic resistance[23]. Moreover, similar to other antibiotics, OTC disturbs microbial communities, resulting in long-term changes in the community characteristics[24].

Bacteriophage therapy, which employs bacteriophages to eliminate bacterial pathogens, is an encouraging substitute for antibiotics[25]. Bacteriophages (phages) represent a distinct group of viruses that selectively infect and lyse bacterial cells. Given their inherent specificity, virulent phages are promising therapeutic agents to combat bacterial diseases[26]. Currently, phage therapy is not commonly used to treat human infections on a global

Department of Biotechnology and Food Science, Norwegian University of Science and Technology, 7491 Trondheim, Norway. ✉email: Ingrid.bakke@ntnu.no

level, however it is increasingly used in aquaculture[27]. Due to their high specificity down to the bacterial strain level, phages are assumed to minimally influence bacterial communities. Many studies confirm this assumption (e.g.[28–30]), however, others have found that phage therapy affects the host-associated microbiota, likely through secondary effects due to lysis of the host bacterium[31–33].

A potential target for phage therapy is columnaris disease, which is a major concern in aquaculture, especially for warm-water salmonid species[34,35]. This disease is caused by the bacterium *Flavobacterium columnare*, which infects, amongst others, salmonid fish[4,36,37]. While antibiotics is the common line of treatment, bacteriophage therapy against *F. columnare* in aquaculture systems has been demonstrated with promising results[38–40].

Here, we used a germ-free model of Atlantic salmon yolk sac fry to investigate the protective effect of the commensal microbiota against infection of *F. columnare*. We therefore challenged both germ-free and microbially colonized fish with the pathogen and compared the observed mortalities. We further used phage therapy against *F. columnare* and investigated its effect on the survival of the fish in comparison to treatment with the antibiotic OTC. Furthermore, we examined the impact of phage treatment on the bacterial communities, expecting it to be minimal especially in the absence of the pathogen. For comparison, antibiotic treatment was applied and changes of the bacterial community structure and α-diversity were compared to that observed during phage therapy.

Materials and methods

Designs of challenge experiments

Three experiments were conducted between autumn 2021 and summer 2022 where fish were challenged with a bacterial pathogen. The first experiment was performed with the purpose of determining the optimal temperature for infecting Atlantic salmon yolk sac fry (*Salmo salar*) with the bacterial pathogen *Flavobacterium columnare* strain Fc7. All fish used in Exp. 1 were raised germ-free (see below). Groups of 15 fish were reared in 250 ml cell culture flasks with aerated caps (hereafter referred to as "fish culture flasks") at 5.6 ± 0.4 °C until 5 weeks post hatching (wph; hatching day is when 70% of all eggs have hatched). At 5 wph, the temperature in the fish culture flasks was gradually increased to either 10 or 14 °C over the course of 1 week or were kept at 6 °C (five replicate flasks per temperature and 15 flasks in total). At 6 wph and for each temperature group, three flasks were challenged with *F. columnare* Fc7, one was exposed to the none-infectious fish commensal *Janthinobacterium* sp. 3.108 (as a control to check whether the mere addition of large loads of a bacterial strain is affecting the survival of the fish) and one was kept as a non-challenged control. The mortality in each flask was checked at least two times daily until the experiment was terminated at 10 days post challenge (dpc).

In experiment 2 (Exp. 2), we challenged Atlantic salmon yolk sac fry with *F. columnare* Fc7 at 10 °C and subsequently added bacteriophages or oxytetracycline to the water daily. We aimed to evaluate the impact of phage- and antibiotic treatment on fish survival and the bacterial community of both the rearing water and the fish. By comparing the survival after challenge with *F. columnare* Fc7 between germ-free and microbially colonized fish (i.e., fish that were hatched under germ-free conditions but were then re-colonized by bacteria; see below), we investigated whether the microbiota protected the yolk sac fry. Both germ-free and colonized fish were raised until 5 wph at 5.2 ± 0.4 °C. Over the next week the rearing temperature was increased to 10 °C and at 6 wph the fish were either challenged with *F. columnare* Fc7 (experimental group hereafter referred to as "Fc7") or left unchallenged ("Control"). Next, both challenged and unchallenged flasks were either treated with the antibiotic oxytetracycline ("AB"), the bacteriophage FCL-2 ("Phage") or kept untreated ("None"), which resulted in the following six experimental groups: Control_None, Control_Phage, Control_AB, Fc7_None, Fc7_Phage, Fc7_AB. Each condition was replicated in three replicate flasks yielding a total of 36 fish culture flasks, containing 15 fish each (2 microbial states (germ-free/colonized) × 2 challenge states (Control/Fc7) × 3 treatment (AB/Phage/None) × 3 biological replicates; Fig. 1). The mortality in each flask was assessed regularly until the experiment was terminated at 10 dpc. Additionally, samples were taken from the rearing water and from three fish per flask at 0 dpc (before challenge), 0.5, 2 and 10 dpc for microbiome analysis (see below). Unfortunately, no sequencing data could be obtained for most of the fish samples, and samples from the fish microbiota are therefore not included in this analysis. Further, the bacterial density in the water was quantified by flow cytometry right before and after *F. columnare* was added to the samples (− 0 and + 0 dpc samples) and at 0.5, 1, 2, 4, 6, 8 and 10 dpc.

Experiment 3 had the same setup as Exp. 2, with the difference that the temperature was increased to 14 °C over the course of 1 week at 5 wph (Fig. 1). We increased the temperature in this experiment to 14 °C as we unexpectedly did not see mortality in the fish in Exp. 2. Therefore, Exp. 3 was conducted to investigate whether *F. columnare* Fc7 induced mortality at 14 °C, whether possessing a microbiota protects the fish against mortality and finally to evaluate whether phage- or antibiotic treatments impacts fish viability. No microbiota samples were taken from this experiment.

Fish husbandry

We raised the Atlantic salmon as described by Gomez de la Torre Canny et al.[41]. In brief, Atlantic salmon eggs were received at ca. 78% development from AquaGen AS (Hemne, Norway) and immediately transferred to a dark room at 6 °C. Groups of 100 eggs were placed in a petri dish (13.5 cm Ø) and covered with Salmon Gnotobiotic Medium (SGM; 0.5 mM $MgSO_4$, 0.054 mM KCl, 0.349 mM $CaSO_4$ and 1.143 mM $NaHCO_3$ dissolved in ultrapure water, autoclaved prior to use at 121 °C for 20 min). After 1 day, the eggs were sterilized (see below) and distributed into 250 ml cell culture flasks with a vented cap, containing 100 ml SGM and 17 fish eggs each. The eggs, and later the fish, were reared in these flasks until the end of the experiment. To maintain good water quality, 60% of the rearing water was exchanged 3 times a week. Dead fish were removed as soon as they were observed. For sampling and for terminating the experiment, fish were euthanised by a lethal dose of tricaine (20 mM, 0.2 µm sterile filtered). As the yolk sac fry that was used in this study was not considered as live animals

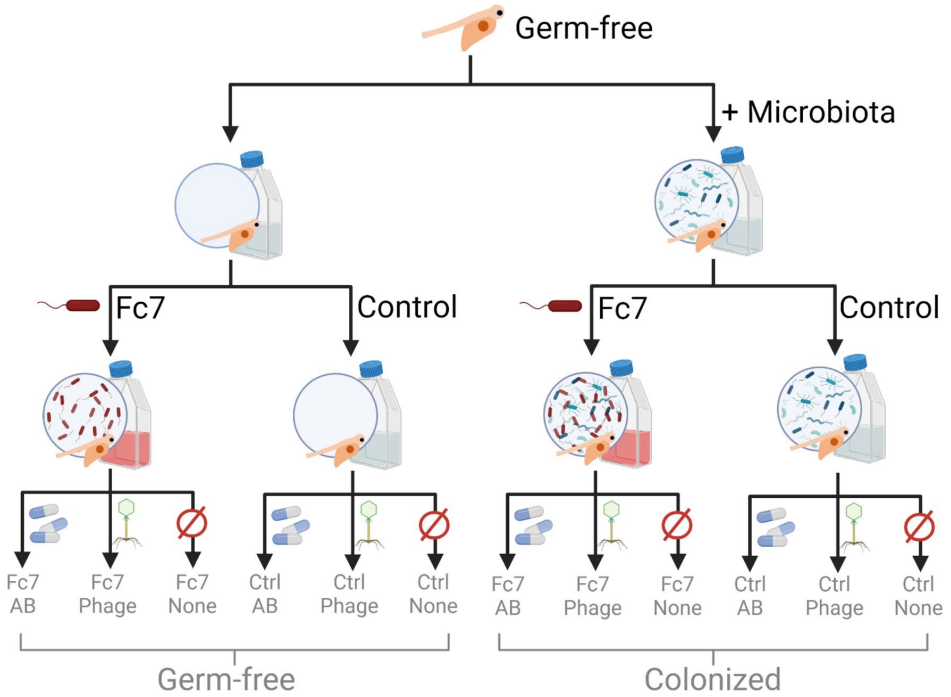


Figure 1. Experimental design of Exp. 2 and Exp. 3, resulting in twelve experimental groups each. Each group consisted of three replicate flasks (36 in total), which each contained 15 fish. Antibiotics and phages were added daily throughout the 10 days following infection at 6 weeks post hatching (wph). Figure created with BioRender.com.

under Norwegian legislation, the experiments conducted in this study had not to be approved by an animal welfare committee.

Sterilization of fish eggs and reintroduction of bacteria (colonization)

The eggs were disinfected as described by Gomez de la Torre Canny and co-workers[41]. First, eggs were surface-sterilized 24 h after arrival at our laboratory by submerging them in an antibiotic cocktail (10 mg l^{-1} Rifampicin, 10 mg l^{-1} Erythromycin, 10 mg l^{-1} Kanamycin, 100 mg l^{-1} Ampicillin, 250 µg l^{-1} Amphotericin B, 150 mg l^{-1} Penicillin and 75 mg l^{-1} Oxolinic acid; antibiotics supplied by Millipore, Sigma-Aldrich, Biowest and Alfa Aesar) for 24 h. Second, groups of 17 eggs were incubated in a Buffodine® solution (FishTech AS) containing 50 mg l^{-1} available iodine for 30 min. They were washed four times with 50 ml SGM and were then placed into a 250 ml cell culture flask containing 100 ml SGM.

To confirm axenity, sterility checks were performed on the hatching day, 1 week before challenge and at the end of the experiment. The sterility check was conducted for each germ-free flask by adding 100 µl rearing water to 3 ml of four different liquid media (Brain Heart Infusion, Glucose Yeast Broth, Sabourad-Dextrose Broth and Nutrient Broth) and a TSB agar plate. The four broths and the plates were incubated at room temperature (RT) for 3 weeks. If growth was observed in either medium, the flask was considered contaminated and removed from the experiment. In addition, water samples were taken and analysed for presence of bacteria using a flow cytometer (Attune NxT, ThermoFisher).

For generating microbially colonized flasks from germ-free flasks, bacteria were reintroduced at 1 wph by adding 60 ml of water from Lake Jonsvatnet (Trondheim, Norway) during the water change. The water was taken from a depth of 50 m in February 2022 (Exp. 2) and in June 2022 (Exp. 3).

F. columnare challenge and treatment of the fish with oxytetracycline or bacteriophage FCL-2

The number of fish was adjusted to 15 fish per flask at 5 wph and, depending on the experiment and experimental group, the temperature was steadily increased to either 9.8 ± 0.3 °C, 14.1 ± 0.3 °C or was kept at 5.6 ± 0.4 °C. For challenging the fish, *F. columnare* Fc7 (kindly provided by David Perez-Pascual and Jean-Marc Ghigo, Institute Pasteur, Paris[4]) was grown in liquid TYES medium (0.5 g l^{-1} $MgSO_4$ * 7 H_2O, 0.2 g l^{-1} $CaCl_2$ * 2 H_2O, 0.4 g l^{-1} yeast extract, 4 g l^{-1} tryptone, 0.5 g l^{-1} D-glucose) at RT and 180 rpm overnight and harvested in late exponential phase at an OD_{600} of approximately 1. The bacterial culture was first spun down at 13,000×*g* for 10 min, then the pellet

was washed with SGM once and finally resuspended in SGM resulting in a concentration of 1×10^9 CFU ml^{-1}. One ml was added per challenged fish culture flask, resulting in a theoretical final concentration of about 10^7 *F. columnare* Fc7 CFUs ml^{-1} in the challenged fish culture flasks.

In Exp. 2 and 3, treatment with either the phage FCL-2 against *F. columnare* Fc7 or the antibiotic oxytetracycline was applied immediately after challenge and daily the next 10 days. For phage treatment we added 1×10^9 PFUs per fish culture flask daily yielding an MOI (multiplicity of infection) of 1 at 0 dpc. For the antibiotic treatment, 4 mg of oxytetracycline (0.4 g/l 0.2 µm sterile filtered stock) was added daily per flask, initially yielding 40 mg l^{-1} per flask at 0 dpc. Nothing was added to the untreated control flasks.

Preparation of phage stock

Phage strain FCL-2[38] against *F. columnare* was kindly provided by Lotta-Riina Sundberg (University of Jyväskylä). Susceptibility of *F. columnare* Fc7 against FCL-2 was confirmed using the soft-agar overlay method. For that, *F. columnare* Fc7 was grown in liquid TYES medium at RT and 180 rpm under aerobic conditions until the exponential phase. Of this culture, 1 ml was added to 3 ml 50 °C warm soft TYES agar (TYES broth containing 7.5 g l^{-1} agar), vortexed and poured out on a TYES agar plate (containing 15 g l^{-1} agar). The plate was incubated at RT for 1 h before 5 µl of the phage stock was added onto the plate. Formation of plaques indicated that *F. columnare* Fc7 was susceptible towards phage strain FCL-2. A phage stock was prepared by harvesting phages from soft TYES agar plates as propagation of the phage in liquid culture was not possible[38]. For that, 100 µl of phage solution (undefined concentration) were mixed with 3 ml soft TYES agar and 1 ml culture of strain Fc7 and was poured on a TYES agar plate, which was incubated at RT overnight. The soft top agar containing the phages was scraped off and suspended in sterile SM buffer (5.8 g l^{-1} NaCl, 50 ml l^{-1} Tris buffer (1 M, pH 7.5), 2 g l^{-1} $MgSO_4$ * 7 H_2O) at a volume ratio of 1:1 soft agar to SM buffer. The mixture was vortexed, spun down at 5000×*g* for 10 min and filtered through a 0.2 µm filter. The phage stock was concentrated by first centrifuging at 22,000×*g* for at least 8 h, before the supernatant was removed and the phage pellet resuspended in a small volume of SM buffer. The titer of the phage stock was determined by spotting out serial dilutions of the stock on soft-agar-overlaid plates and counting plaques.

Flow cytometry analysis

We quantified the bacterial density in the rearing water in Exp. 2 using flow cytometry at nine different timepoints: Before and after *F. columnare* Fc7 was added to the samples (− 0 and + 0 dpc samples) and at 0.5, 1, 2, 4, 6, 8 and 10 dpc. We sampled 1 ml water per flask and sampling time. The water samples were fixated in 0.1% glutaraldehyde for 15 min before they were snap-frozen in liquid nitrogen and stored at − 80 °C until data acquisition. Samples were diluted in 0.2 µm-filtered phosphate-buffered saline to obtain stable sample aquation and dilute background noise present in the sample. The water samples were stained with the DNA-binding fluorescent dye SYBR green I (Invitrogen, final concentration 2 ×), vortexed and incubated for 15 min in the dark at 37 °C. Stained samples were vortexed and 160 µl was sampled at a 100 µl min^{-1} flowrate. A zip clean of the instrument (Attune NxT Flow Cytometer) was performed between approximately every 6th sample. Data were collected using the blue laser (488 nm) with detection in BL1 (530/30 nm) and BL3 (695/40 nm) using a BL1 threshold of 1500–3000 (depending on sample). Instrument voltages were as follows; FSC 320V, SSC 260V, BL1 320 V and BL3 350V. Filtered PBS and 0.2 µm-filtered fish rearing water were used as negative controls and a pure culture of strain Fc7 as positive control. Identification of bacterial populations was achieved by comparing 0.2 µm-filtered fish rearing water with unfiltered rearing water. All samples were gated in the same way with some minor modifications to each sample gate to ensure high quality data (https://figshare.com/articles/dataset/fsc_files_Fc7_Salmon/21518922/1).

Sampling for characterization of the water microbiota in Exp. 2

Water samples for microbiome analysis were taken only for Exp. 2. Samples were taken at four timepoints and only from colonized flasks: Before challenge with *F. columnare* Fc7 (0 dpc) and at 0.5, 2 and 10 dpc. For each sampling time we collected the rearing water bacterial community by filtering 10 ml rearing water through a 0.2 µm polycarbonate filter (Osmonics, 25 mm diameter). The filter was cut in pieces and transferred to a BeadBashing tube of the ZymoBIOMICS™ 96 MagBead DNA kit (Zymo), snap-frozen in liquid nitrogen and stored at − 80 °C. In total 72 water samples were taken.

DNA extraction

DNA was extracted from all water samples using the ZymoBIOMICS™ 96 MagBead DNA kit (Zymo) and a KingFisher Flex instrument. The filters representing the water samples were cut in smaller pieces using a sterile scalpel prior to extraction of DNA. The filters were homogenized by adding 750 µl (450 µl for 0 dpc samples) lysis buffer from the DNA extraction kit and run two cycles à 30 s at 5500 rpm in a Precellys 24 (Bertin Technologies). The homogenized samples were spun down at 13,000×*g* for 10 min. For samples from 0 dpc, we used 450 µl of the homogenate supernatant, whereas 200 µl were used from the 0.5, 2 and 10 dpc samples. DNA was extracted following the manufacturer's protocol except for final elution in 100 µl DNAse-free water (50 µl for 0 dpc samples). The extracted DNA was stored at − 20 °C until library preparation. Additionally, two negative controls using only lysis buffer as input for DNA extraction were prepared.

16S rRNA amplicon sequencing

The v3 + v4 region of the 16S rRNA gene was amplified using the PCR primers with Illumina tag Ill-341F (5′-TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGNNNN**CCTACGGGNGGCWGCAG**-3′) and Ill-805R (5′-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGNNNN**GACTACNVGGGTATCTAAKCC**-3′), with

the target sequences shown in bold[42,43]. PCR was conducted in 25 µl reaction volumes, where each reaction contained 0.15 µM of each primer, 0.25 µM of each dNTP as well as 0.4 U Phusion hot start polymerase and the respective buffer from Thermo Scientific. For samples from 0 dpc, 1 µl of a 1:100 dilution of the DNA extract was used as template, whereas 1 µl undiluted DNA extract was used as template for the other samples. The cycling conditions were as follows: an initial denaturation step at 98 °C for 60 s followed by 33 cycles of 98 °C for 15 s, 55 °C for 20 s and 72 °C for 20 s. The final elongation step was 72 °C for 5 min before the samples were cooled to 10 °C. PCR products were evaluated by electrophoresis on a 1% agarose gel containing 50 µM GelRed (Biotium) for 1 h at 110 V. The samples were normalized using Sequal Prep™ Normalization plates (96 wells, Invitrogen) and indexed using the Nextera® XT Index Kit v2 Sets A and D. For indexing, 2.5 µl normalized PCR product was used as template with 2.5 µl of each indexing primer, 0.2 µM of each dNTP, 0.4 U Phusion hot start polymerase and its buffer from Thermo Scientific (total volume 25 µl). The same cycling program as above was run with 12 cycles. The indexed PCR products were again normalized by loading 15 µl indexed PCR products onto Sequal Prep™ Normalization plates and were pooled and concentrated using a Amicon® Ultra 0.5 ml centrifugal filter (30K membrane, Merck Millipore). A NanoDrop™ One Microvolume Spectrophotometer (Thermo Scientific™) was used to evaluate the quality and quantity of the DNA. The samples were sent to the Norwegian Sequencing Center using one run on a MiSeq v3 instrument with 300 paired ends. The sequencing data was deposited at the European Nucleotide Archive (ERS14896569-ERS14896640).

Analysis of the Illumina sequencing data

The USEARCH pipeline (v.11)[44] was used to process the data obtained from Illumina sequencing. The sequencing pairs were merged, and primer sequences trimmed off using the Fastq_mergepairs command with a minimal length of 390 bp. The merged sequences were quality-filtered using the Fastq-filter function with the default error threshold value of 1. The reads were pooled, dereplicated and singleton reads removed. Amplicon sequence variants (ASVs) were generated using the Unoise2 command[45] with the default minimum abundance threshold of 8 reads in the total dataset. Taxonomical assignment of the ASVs was achieved using the Sintax command[46] with a confidence threshold of 0.8 and the RDP reference dataset v. 18[47]. Reads classified as eukaryotes or chloroplasts were removed from the data set. A few ASVs that were highly abundant in negative controls for the DNA extraction or the phage stock, but less abundant in the samples, were considered to represent contaminating DNA and were removed from the data set. One sample (replicate flask 2, 0.5 dpc, Control-Phage group) was removed from the dataset as it had extremely poor sequencing efficiency compared to all other samples. In the final ASV table, samples had 108,719 ± 20,015 reads on average and were normalized by scaling to 74,545 reads per sample. All analyses were performed using the normalized ASV table. By using the BLAST algorithm[48] the 16S rDNA sequence of *F. columnare* Fc7 (Supp. Fig. 1) was compared to all ASVs of the dataset and we identified ASV3 to correspond to *F. columnare* Fc7.

Statistical analysis

All statistical analyses were performed in R (v. 4.1.3[49]) using RStudio (2022.07.1) and the packages Phyloseq (v. 1.38.0[50]), Vegan (v. 2.6.2[51]), ggplot2 (v. 3.3.6), dplyr (v. 1.0.9), reshape2 (v. 1.4.4), genefilter (1.76.0), DECIPHER (v. 2.22.0) and ggh4x (v. 0.2.2.9000). The *renyi* function of Vegan was used to calculate the α-diversities of samples as Hill's diversity numbers[52,53]. Ordination by principal coordinate analysis (PcoA) was performed using the *ordinate* function from phyloseq for Bray–Curtis dissimilarities, if not stated otherwise. Phylogenetic trees were generated using the phangorn package in R (v. 2.9.0[54]) by first constructing a neighbour-joining tree and then fitting a GTR + G + I model to it. PERMANOVA analyses[55] based on Bray–Curtis dissimilarities (if not stated otherwise) were performed using the *adonis*2 function from vegan by running it in 100 iterations with 999 permutations each and the mean p-value of the 100 iterations was reported (mathematically lowest possible p-value = 0.001). For statistical univariate data (e.g. α-diversity indices or abundance of certain ASVs), the data was checked for normality using the Shapiro–Wilk test (*hapiro.test* function). When the data were found to be normally distributed, a Welch's *t*-test (*t.test* function) was used for data with two groups and ANOVA (*aov* function) for data with three groups. A Mann–Whitney *U* test (*wilcox.test* function) or Kruskal–Wallis test (*kruskal.test* function) was performed for these purposes on non-normally distributed data. Significant ANOVA or Kruskal–Wallis tests were followed by a Bonferroni-corrected Dunn's test (*dunnTest* function). A significance level of $\alpha < 0.05$ was used for all analyses. All box plots are presented as median and upper and lower quartile as box, whiskers include all samples except for outliers. Kaplan–Meier survival analysis was performed using the survival (v. 3.3.1) and survminer (v. 0.4.9) packages in R.

Results

Challenging Atlantic salmon yolk sac fry with *Flavobacterium columnare* Fc7 at different temperatures

In the first experiment (Exp. 1), we examined whether the bacterial pathogen *Flavobacterium columnare* Fc7 induced a lethal infection in germ-free Atlantic salmon yolk sac fry at 6, 10 or 14 °C. Both at 10 and 14 °C, all fish died within 60 h and 48 h, respectively, in all three replicate flasks that had been challenged with *F. columnare Fc7* (data not shown). No mortality was observed in flasks challenged at 6 °C or in any of the control flasks that had not been added *F. columnare* Fc7. The exception was one dead fish in a flask that was exposed to the bacterial commensal *Janthinobacterium* sp. 3.108 at 14 °C. We therefore concluded that *F. columnare* Fc7 induced mortality in germ-free Atlantic salmon yolk sac fry at both 10 and 14 °C.

F. columnare Fc7 challenge and phage treatment at 10 °C in Exp. 2

As yolk sac fry thrive best at lower temperatures, we performed the second challenge experiment (Exp. 2) at 10 °C. We challenged both germ-free and microbially colonized fish with *F. columnare* Fc7 to assess whether the presence of a microbiota protects the fish against lethal infections. By consecutively treating them with either the bacteriophage FCL-2 or the antibiotic oxytetracycline (OTC) we wanted to examine whether bacteriophage treatment can be used to protect the fish against infection and further assessed the effect of phage- and antibiotic treatment on the water and fish microbiota by 16S rRNA gene sequencing (see Fig. 1 and "Materials and methods").

Effect of phage therapy on fish survival at 10 °C

Unexpectedly, *F. columnare* Fc7 did not induce mortality in neither the germ-free nor the colonized fish in Exp. 2 as no mortality was observed throughout the whole experiment in all flasks. Thus, we could not reproduce the observed mortality of Exp. 1. Therefore, we could not draw a conclusion whether phage treatment or presence of a microbiota protected the fish against lethal infections in Exp. 2.

Effect of phage FCL-2 on the relative amounts of *F. columnare* Fc7 in the water microbiota

We sampled rearing water and fish from Exp. 2 at 0, 0.5, 2 and 10 days post challenge (dpc) in order to investigate the impact of phage- and antibiotic treatment on the water and fish microbiota by 16S rRNA gene sequencing. Unfortunately, the amplification of the bacterial 16S rRNA gene from most fish samples was unsuccessful, and therefore, only the bacterial community of the water samples is presented.

We first investigated whether phage therapy reduced the relative abundance of the pathogen in the rearing water. We did not observe the ASV representing *F. columnare* Fc7 (ASV3) in the water prior to the *F. columnare* Fc7 challenge (0 dpc samples; Fig. 2). In samples collected at 0.5 dpc, the relative average abundance of ASV3 in the bacterial communities of the water was 31 ± 7% in the flasks to which strain Fc7 was added (Fc7-flasks; Fig. 2). There was no difference in the relative abundance of ASV3 at 0.5 dpc among the Fc7-flasks (Kruskal–Wallis test, $p = 0.433$), indicating that similar amounts of *F. columnare* Fc7 were added to each flask. The relative abundance of ASV3 decreased in both the Fc7_None and the Fc7_Phage flasks. This decrease was on average 76-fold from 0.5 to 10 dpc in Fc7_Phage flasks but only ninefold in Fc7_None flasks (Fig. 2). Due to the limited numbers of replicates, no statistical test could be conducted to confirm whether the difference between treatments were significant. Further, the relative abundance of *F. columnare* Fc7 decreased faster in most Fc7_Phage flasks than in the Fc7_None flasks (Fig. 2), indicating that phage therapy successfully reduced the populations of *F. columnare* Fc7 in the flasks.

Unexpectedly, the antibiotic treatment did not affect the relative abundance of ASV3 in the water community profiles (Fig. 2). Moreover, the relative abundance of ASV3 in Fc7_AB flasks was significantly higher than in Fc7_None and Fc7_Phage flasks at 10 dpc (Kruskal–Wallis test, $p = 0.049$). This was surprising, as preliminary tests confirmed susceptibility of *F. columnare* Fc7 to OTC on TYES agar plates (data not shown).

In conclusion, the relative abundance of *F. columnare* Fc7 decreased in both Fc7-None and Fc7-Phage flasks, with a slightly stronger decrease in Fc7-Phage flasks, while no decrease was observed in Fc7-AB flasks.

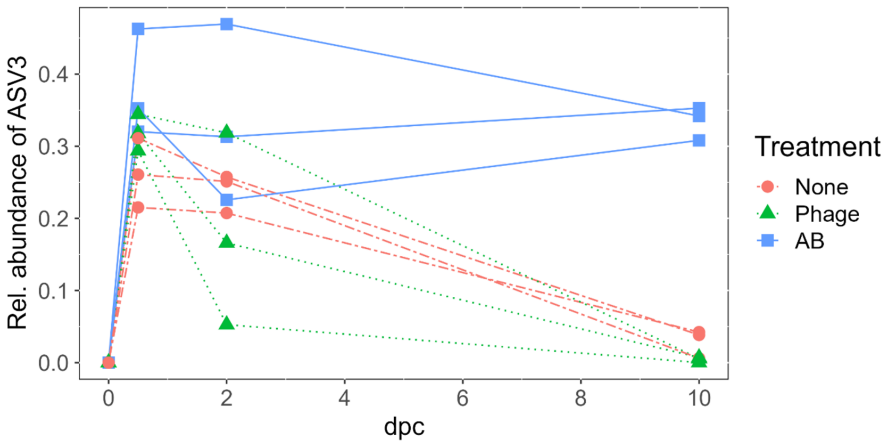


Figure 2. The relative abundance of ASV3 in the water microbiota of all Fc7-flasks of Exp. 2. ASV3 was identified to represent *F. columnare* Fc7. Each line connects the observations within one rearing flask, and colours indicate the treatment application (None = control treatment, Phage = FCL-2 addition, AB = oxytetracycline addition; dpc = days post challenge).

Impact of *F. columnare* Fc7 challenge on the rearing water microbiota

The water microbiota differed significantly between all flasks already at 0 dpc, prior to challenge with *F. columnare* Fc7 even though all flasks were treated identically until the challenge with *F. columnare* Fc7 (Figs. 3 and 4; PERMANOVA: p-value = 0.001). These differences were therefore most likely a result of individual bacterial community succession in each flask during the 6 weeks before the challenge and not of the experimental groups.

The microbiota in the Control-flasks (no *F. columnare* Fc7 added) did not change significantly during the experiment (PERMANOVA, $p > 0.580$ for all three comparisons; Fig. 3). However, in the Fc7-flasks, the water microbiota had changed 12 h after addition of strain Fc7, mainly due to an increase in the relative abundance of ASV3 (Fig. 3 and Supp. Fig. 2). Interestingly, PCoA indicated that 10 days after addition of *F. columnare* Fc7, the water microbiota in Fc7-Phage and Fc7-None flasks returned to their composition prior to the addition of *F. columnare* Fc7 at 0 dpc (Fig. 3). This recovery was, however, not observed in Fc7_AB flasks (Fig. 3), as the relative abundance of ASV3 remained high (Fig. 2 and Supp. Fig. 2).

These temporal effects in the Fc7-flasks were also observed when ASV3 was removed from the dataset prior to PCoA ordination (Supp. Fig. 3) and when ordinations were based on presence-absence data (Sørensen-Dice dissimilarity, Supp. Fig. 4). These findings show that addition of *F. columnare* Fc7 affected the water microbiota in the challenged flasks.

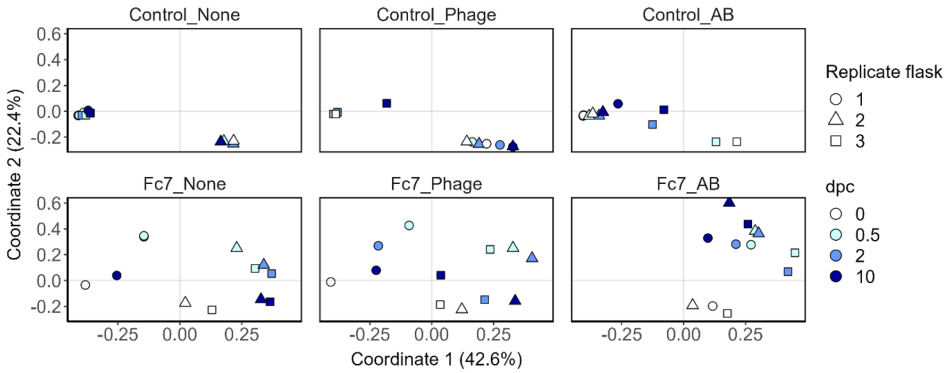


Figure 3. PCoA of the Bray–Curtis dissimilarities of all samples from Exp. 2. Sampling timepoints are represented by different colours, whereas biological replicate flasks are indicated by different shapes. All panels are from the same PCoA but were facetted into the different treatment groups.

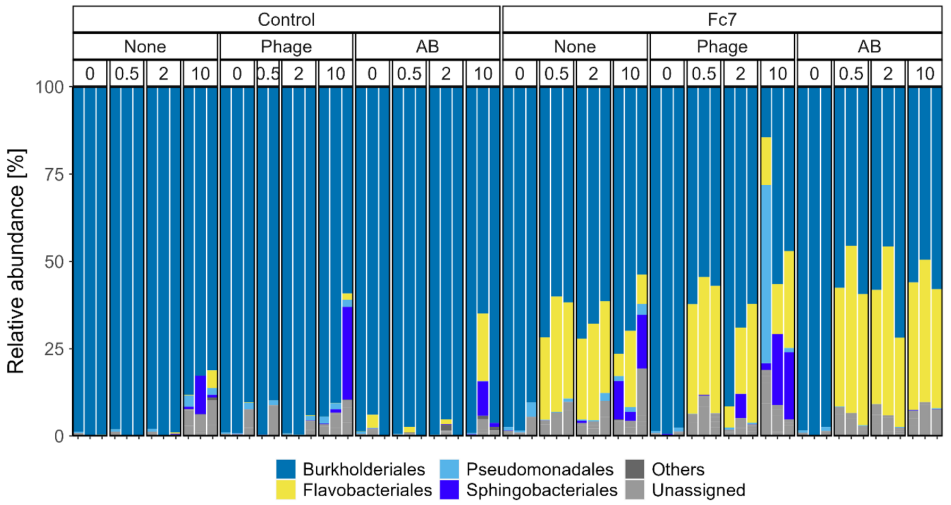


Figure 4. Bacterial community composition of each water sample from Exp. 2, shown at the order level. Orders that are not appearing with at least 5% in at least one sample are summarized as "others".

Effect of phage therapy and antibiotic treatment on the water microbiota

To evaluate the effect of the treatments on the bacterial communities we evaluated the community composition change over time, the α-diversity and the bacterial density.

We therefore compared the bacterial community in the None-flasks with Phage- or AB-flasks over the course of the experiment to assess the effect of phage- and antibiotic treatment both in the presence and absence of *F. columnare* Fc7. Because the bacterial communities differed between replicate flasks from the same experimental group (Supp. Fig. 5) we investigated the temporal change within each individual flask. For each flask, we calculated the Bray–Curtis similarity between each sampling timepoint and the start of the experiment (0 dpc). For Control-flasks, the change in Bray–Curtis similarity over time was comparable between treatments (Fig. 5). This indicated that in the absence of *F. columnare* Fc7, neither the antibiotic- nor the phage treatment influenced the water bacterial community. For the Fc7-flasks, the communities in Fc7_Phage and Fc7_AB flasks changed more over time than Fc7_None flasks (Fig. 5). Therefore, in the presence of the pathogen, both the antibiotic and the phage treatment appeared to influence the bacterial water communities compared to Fc7_None flasks.

Hill's diversity of order 1 (^{1}D; exponential Shannon index) was not significantly affected by the phage treatment (Supp. Fig. 6), as no significant difference was observed when comparing the change in ^{1}D from 0 to 10 dpc for each flask between the different treatments (ANOVA, $p = 0.532$ and 0.592 for Control- and Fc7-flasks, respectively). Highly unexpectedly, AB treatment increased the α-diversity in the Control-flasks by a factor of 2 from 0 to 10 dpc, whereas this was not observed in Control-None flasks (Supp. Fig. 6). Lastly, phage treatment did not decrease the bacterial density when Fc7 was absent (Supp. Fig. 7).

These findings indicate that phage therapy did not influence the bacterial communities in the absence of the phage's host, *F. columnare* Fc7, whereas changes in the microbiota were observed in the presence of the phage's host. Phage therapy did further not significantly affect the α-diversity. Unexpectedly, also the antibiotic oxytetracycline did not disturb the microbiota in the water in the absence of *F. columnare* Fc7.

F. columnare Fc7 challenge and phage treatment at 14 °C in Exp. 3

As no mortality was induced by *F. columnare* Fc7 at 10 °C in Exp. 2, we conducted a third experiment to investigate the protective effect of possessing a microbiota and to determine whether phage and antibiotic treatment reduced mortality in infected fish. The experimental setup was similar between Exp. 2 and 3, with the exception that the temperature was increased to 14 °C to increase infectivity *of F. columnare* Fc7[35].

All germ-free fish in Fc7_None flasks died after challenge with *F. columnare* Fc7. In contrast, mortality was only observed in one flask in colonized Fc7_None flasks (Fig. 6). Thus, the survival was significantly higher when fish possessed a microbiota (Kaplan–Meier survival analysis, $p < 0.0001$). This suggests that the fish microbiota protects the fish against lethal bacterial infection.

As there was no mortality in neither phage- nor the antibiotic-treated germ-free flasks, we concluded that both phage therapy and antibiotic treatment protected the fish against lethal infections (Fig. 6). However, due to high survival in colonized fish we could not conclude on the effect of phage therapy in colonized fish. We further

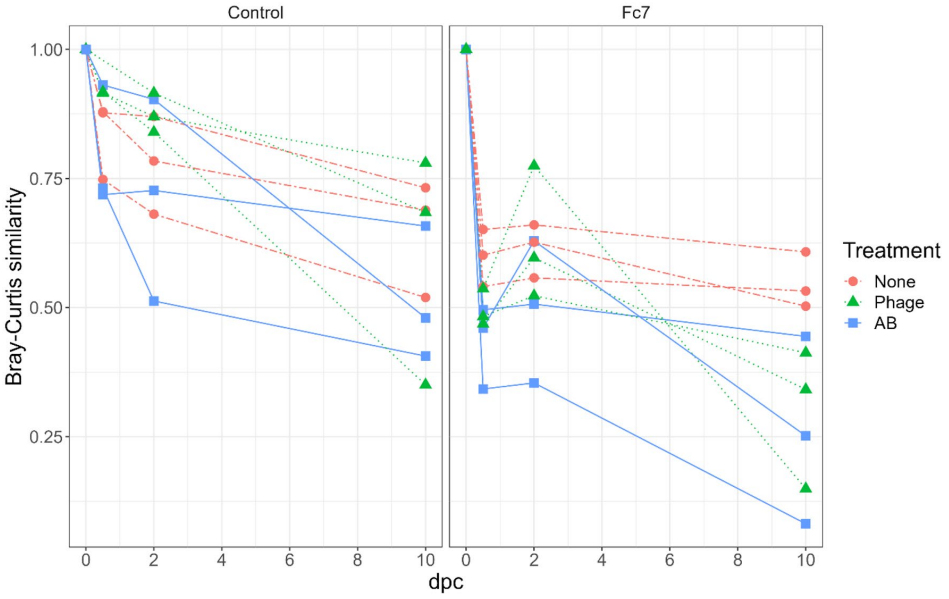


Figure 5. Bray–Curtis similarities within each replicate flask comparing the water bacterial community from each timepoint to the community at 0 dpc. Each line represents one rearing flask, and colours indicate the treatment application (None = control treatment, Phage = FCL-2 addition, AB = oxytetracycline addition).

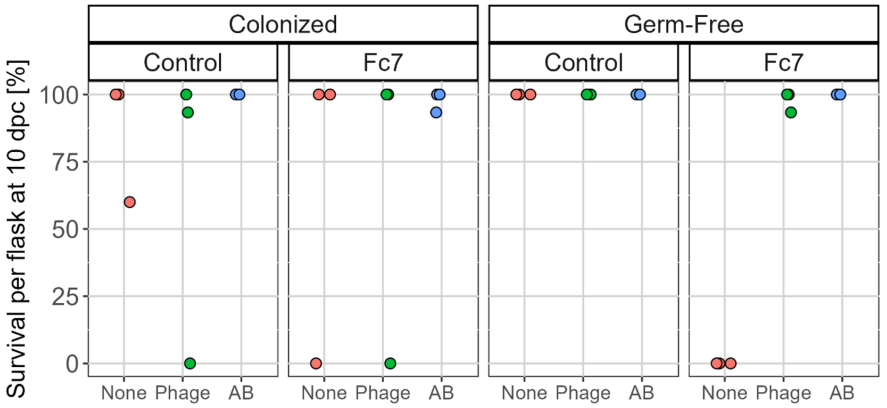


Figure 6. Fish survival at 10 dpc in Exp. 3. Each dot represents survival in one replicate flask (three replicate flasks per group).

observed unexpected mortality in colonized Control-None and Control-Phage flasks to which no pathogen was added. These observations are addressed in the Discussion part.

We therefore concluded that the presence of a microbiota protected the yolk sac fry against lethal infection with *F. columnare* Fc7, and also that both the phage- and the antibiotic treatment protected the fish.

Discussion

We urgently need alternatives to antibiotics as antibiotic resistance genes spread, the development of novel antibiotics is slow and the ones we use have harmful side effects on our microbiome[10,56]. One alternative is phage therapy, where a virulent bacteriophage is used to lyse, and thus kill, a specific pathogenic bacterium[26]. Clinical trials on humans with phage therapy were shown to be challenging due to administrational hurdles and due to a lack of information on how phages affect the human body and its microbiome[57,58].

In aquaculture, several commercial phage-products are already available[27]. The application of phage therapy often increases survival in infection experiments (e.g.[59–61]), although sometimes studies find no benefit from the therapy[62]. While it is generally accepted that phage therapy may alleviate aquatic disease outbreaks, the effect that phage therapy has on the microbiota is poorly documented. In this study, we therefore aimed to investigate the impact of phage therapy on the rearing water of the commercially important species Atlantic salmon.

In order to test phage therapy, we first had to establish a challenge protocol to infect the fish with a bacterial pathogen, and we chose to use *F. columnare* for this purpose. In Exp. 1, we found that *F. columnare* rapidly induced mortality in Atlantic salmon yolk sac fry at both 10 and 14 °C. This observation could be replicated at 14 °C in Exp. 3 but not at 10 °C in Exp. 2. It was puzzling that we observed mortality in germ-free fish at 10 °C in Exp. 1 but not in Exp. 2 as the experimental conditions were identical. Differences between egg batches might have contributed, although it is impossible to investigate this in retrospect. Because *F. columnare* primarily infects fish at warmer temperature, we expected less infection success at lower temperature[35,36]. Still, it was interesting that infection induced mortality abruptly, with only a 4 °C difference between the experiments.

After establishing a challenge protocol, we first investigated whether the microbiota in the fish has a protective effect during the bacterial challenge and therefore compared fish survival in germ-free and colonized fish. All germ-free fish infected with *F. columnare* at 14 °C died within three days. In contrast, for the colonized fish, mortality was only observed in one of the three replicate flasks. Thus, the microbiota of the salmon protects the fish. Such a protective effect of the microbiota is generally assumed to be present in animals, however, studies investigating it in fish are scarce so far[4,5]. It was therefore important to confirm this protective effect in fish, after it had been shown previously in zebrafish[5] and rainbow trout[4].

The high mortality in untreated germ-free fish, and the fact that we saw no mortality in phage-treated germ-free fish, showed that bacteriophage therapy was successful in increasing survival during infection with *F. columnare*. The same phage that we used (FCL-2) has also been successfully used earlier in rainbow trout and zebrafish, where elevated survival was observed[38,40]. Our findings therefore confirm the potential of using this phage in aquaculture against *F. columnare* infections.

Unfortunately, we were not able to draw a definite conclusion about the success of phage therapy in the colonized fish, as survival in the colonized Fc7-flasks was generally very high. Unexpectedly, we occasionally observed that fish suddenly died, even in flasks that were unchallenged with *F. columnare* Fc7. This was likely caused by instabilities in our experimental system at 14 °C: at higher temperature, increased bacterial growth and metabolic activity may deplete oxygen, causing subsequent fish death. This temperature-dependent mortality has been observed previously in similar experimental setups in our research group (data not shown). This explanation seems plausible, as no mortality was observed when OTC was added to the water. OTC decreases metabolic activity by inhibiting protein synthesis. Thus, the bacterial oxygen consumption could not increase when OTC was present in the water, and the fish did not die.

In order to assess the effect of phage therapy on the water bacterial community, we sampled the water microbiota in Exp. 2. We originally aimed to investigate both the fish and the water microbiota. However, for over 70% of the fish samples, we could not generate 16S rRNA gene amplicons and, consequently, all fish samples were discarded from further analysis. Nevertheless, it is relevant to investigate the effect of phage treatment on the water bacterial community, as the water influences the fish microbiota[63–66]. First, we inspected whether phage treatment reduced the relative abundance of *F. columnare* in the water. Here, we expected that the phage strongly reduced the relative abundance of the pathogen due to lysis of the cells. However, the relative abundance of *F. columnare* Fc7 in the water microbiota was quickly reduced, also in flasks that were not treated after the challenge. This shows that *F. columnare* was not able to persist in the rearing flasks under the experimental conditions used in this study. *F. columnare* Fc7 might not be metabolically active at 10 °C[35], and this might decrease the effect of the phage treatment[67,68]. Still, we observed a faster and more pronounced decrease of the relative *F. columnare* Fc7 abundance due to addition of phage FCL-2, which indicates that the phages were lysing the bacterial cells.

Phage therapy has been proposed to exert minor effects on the microbiota because the treatment only targets one specific population[69]. When the phage's target bacterium is absent, no strong effects on the microbiota are expected[70–73]. Concordantly, we observed no significant changes in the community composition, α-diversity or bacterial density due to the FCL-2 treatment in the absence of *F. columnare*. When the phage's target is present, however, the removal of the target bacterial population can affect the community properties[31,33]. These effects depend on the interaction network of the target population within the community and its abundance. *F. columnare* Fc7 was not present in the water communities prior to the Fc7 challenge and is therefore most likely not part of the interaction network in the microbial communities. However, we added *F. columnare* Fc7 in high abundances to the Fc7_Phage flasks. It was therefore not surprising to observe changes in the bacterial compositions in these flasks. These small changes might be due to lysis of Fc7-cells which liberates nutrients to the other community members[31]. It would have been interesting to investigate the effect of phage therapy on the microbiota during a natural outbreak of columnaris disease. In such a scenario the removal of the pathogenic population might result in stronger downstream effects due to disruption of the bacterial interaction-network[31]. Even though slight changes in the bacterial community occurred, no effect on the α-diversity or absolute abundance of the cells was observed. With this, our study adds to the growing body of evidence, that phage therapy does not cause negative side-effects on the bacterial community[30,59,74–76].

Surprisingly, antibiotic treatment with OTC did not affect the microbiota as we would have expected it. The antibiotic treatment did not reduce the relative abundance of *F. columnare* Fc7 and did further not distort the microbiota in the flasks or decrease α-diversity. This was unexpected, as OTC is a widely-used broad-spectrum antibiotic that has been shown to elicit strong disturbances onto the microbiota and to decrease the bacterial richness[19,20,77–79]. It is also generally accepted that antibiotic treatment in general has a disruptive effect on the microbiota[80–82]. Nevertheless, some studies find that OTC does not always disturb the microbiota or decrease the richness[83,84]. A reason for why we did not see stronger effects on the bacterial community and the abundance of *F. columnare* Fc7 could be that OTC's function was somehow impaired. However, its efficacy against *F. columnare* Fc7 was tested before, during and after the experiment and a fresh AB stock was prepared at 5 dpc in order to avoid degradation of OTC over time. Furthermore, the higher survival of the antibiotic-treated fish in Exp. 3 indicated that OTC was active and functioning. Another explanation could be that the bacteria in the water are slow growing at 10 °C, which could delay observable effects by the bacteriostatic OTC. Since we unexpectedly did not observe disturbances induced by OTC, it would had been beneficial to use a bactericidal antibiotic for disturbing the bacterial communities, which would have allowed better comparison of the effects of antibiotics and phage therapy on the fish microbiota.

In conclusion we showed that phage therapy protected Atlantic salmon yolk sac fry against infection with *F. columnare*, without disturbing the microbiota of the water. While our work was conducted in Atlantic salmon, we think that our findings can also be useful for discussing usage of phage therapy in a context outside of aquaculture.

Data availability

All flow cytometry files used for bacterial density quantification are available through https://figshare.com/projects/fsc_files/152463. The Illumina sequencing reads are available through the European Nucleotide Archive (ERS14896569-ERS14896640).

Received: 20 June 2023; Accepted: 14 October 2023
Published online: 06 November 2023

References

1. Shreiner, A. B., Kao, J. Y. & Young, V. B. The gut microbiome in health and in disease. *Curr. Opin. Gastroenterol.* **31**, 69–75. https://doi.org/10.1097/Mog.0000000000000139 (2015).
2. Selber-Hnatiw, S. *et al.* Human gut microbiota: Toward an ecology of disease. *Front. Microbiol.* **8**, 1265. https://doi.org/10.3389/fmicb.2017.017E5 (2017).
3. Becattini, S. *et al.* Commensal microbes provide first line defense against *Listeria monocytogenes* infection. *J. Exp. Med.* **214**, 1973–1989. https://doi.org/10.1084/jem.20170495 (2017).
4. Pérez-Pascual, D. *et al.* Gnotobiotic rainbow trout (*Oncorhynchus mykiss*) model reveals endogenous bacteria that protect against *Flavobacterium columnare* infection. *PLoS Pathog.* **17**, e1009302. https://doi.org/10.1371/journal.ppat.1009302 (2021).
5. Stressmann, F. A. *et al.* Mining zebrafish microbiota reveals key community-level resistance against fish pathogen infection. *ISME J.* **15**, 702–719. https://doi.org/10.1038/s41396-020-00807-8 (2021).
6. Kissoyan, K. A. B. *et al.* Natural *C. elegans* microbiota protects against infection via production of a cyclic lipopeptide of the Viscosin group. *Curr. Biol.* **29**, 1030. https://doi.org/10.1016/j.cub.2019.01.050 (2019).

7. Dillon, R. J., Vennard, C. T., Buckling, A. & Charnley, A. K. Diversity of locust gut bacteria protects against pathogen invasion. *Ecol. Lett.* **8**, 1291–1298. https://doi.org/10.1111/j.1461-0248.2005.00828.x (2005).
8. Rosas-Plaza, S. *et al.* Human gut microbiome across different lifestyles: From hunter-gatherers to urban populations. *Front. Microbiol.* **13**, 843170. https://doi.org/10.3389/fmicb.2022.843170 (2022).
9. Panwar, R. B., Sequeira, R. P. & Clarke, T. B. Microbiota-mediated protection against antibiotic-resistant pathogens. *Genes Immun.* **22**, 255–267. https://doi.org/10.1038/s41435-021-00129-5 (2021).
10. Blaser, M. J. Antibiotic use and its consequences for the normal microbiome. *Science* **352**, 544–545. https://doi.org/10.1126/science.aad9358 (2016).
11. Theriot, C. M. *et al.* Antibiotic-induced shifts in the mouse gut microbiome and metabolome increase susceptibility to *Clostridium difficile* infection. *Nat. Commun.* **5**, 3114. https://doi.org/10.1038/ncomms4114 (2014).
12. Watkins, R. R. & Bonomo, R. A. Overview: Global and local impact of antibiotic resistance. *Infect. Dis. Clin. N. Am.* **30**, 313. https://doi.org/10.1016/j.idc.2016.02.001 (2016).
13. European Centre for Disease Prevention and Control (ECDC); European Food Safety Authority (EFSA); European Medicines Agency (EMA). Third joint inter-agency report on integrated analysis of consumption of antimicrobial agents and occurrence of antimicrobial resistance in bacteria from humans and food-producing animals in the EU/EEA. *EFSA Journal* **19**, e06712. https://doi.org/10.2903/j.efsa.2021.6712 (2021).
14. Chopra, I. & Roberts, M. Tetracycline antibiotics: Mode of action, applications, molecular biology, and epidemiology of bacterial resistance. *Microbiol. Mol. Biol. Rev.* **65**, 232–260. https://doi.org/10.1128/MMBR.65.2.232-260.2001 (2001) **(second page, table of contents)**.
15. Brodersen, D. E. *et al.* The structural basis for the action of the antibiotics tetracycline, pactamycin, and hygromycin B on the 30S ribosomal subunit. *Cell* **103**, 1143–1154. https://doi.org/10.1016/s0092-8674(00)00216-6 (2000).
16. Monteiro, S. H. *et al.* Study of spatial and temporal distribution of antimicrobial in water and sediments from caging fish farms by on-line SPE-LC-MS/MS. *J. Environ. Sci. Health B* **51**, 634–643. https://doi.org/10.1080/03601234.2016.1181917 (2016).
17. Capone, D. G., Weston, D. P., Miller, V. & Shoemaker, C. Antibacterial residues in marine sediments and invertebrates following chemotherapy in aquaculture. *Aquaculture* **145**, 55–75. https://doi.org/10.1016/S0044-8486(96)01330-0 (1996).
18. Coyne, R., Hiney, M. & Smith, P. Transient presence of oxytetracycline in blue mussels (*Mytilus edulis*) following its therapeutic use at a marine Atlantic salmon farm. *Aquaculture* **149**, 175–181. https://doi.org/10.1016/S0044-8486(96)01446-9 (1997).
19. Qian, M. R. *et al.* Sub-chronic exposure to antibiotics doxycycline, oxytetracycline or florfenicol impacts gut barrier and induces gut microbiota dysbiosis in adult zebrafish (*Daino rerio*). *Ecotoxicol. Environ. Saf.* **221**, 112464. https://doi.org/10.1016/j.ecoenv.2021.112464 (2021).
20. Zhou, L. *et al.* Environmental concentrations of antibiotics impair zebrafish gut health. *Environ. Pollut.* **235**, 245–254. https://doi.org/10.1016/j.envpol.2017.12.073 (2018).
21. Rodrigues, S., Antunes, S. C., Correia, A. T. & Nunes, B. Rainbow trout (*Oncorhynchus mykiss*) pro-oxidant and genotoxic responses following acute and chronic exposure to the antibiotic oxytetracycline. *Ecotoxicology* **26**, 104–117. https://doi.org/10.1007/s10646-016-1746-3 (2017).
22. Kayani, M. U. R. *et al.* Environmental concentrations of antibiotics alter the zebrafish gut microbiome structure and potential functions. *Environ. Pollut.* **278**, 116760. https://doi.org/10.1016/j.envpol.2021.116760 (2021).
23. Tian, Z. *et al.* Minimum influent concentrations of oxytetracycline, streptomycin and spiramycin in selecting antibiotic resistance in biofilm type wastewater treatment systems. *Sci. Total Environ.* **720**, 137531. https://doi.org/10.1016/j.scitotenv.2020.137531 (2020).
24. Jakobsson, H. E. *et al.* Short-term antibiotic treatment has differing long-term impacts on the human throat and gut microbiome. *PLoS One* **5**, e9836. https://doi.org/10.1371/journal.pone.0009836 (2010).
25. Salmond, G. P. C. & Fineran, P. C. A century of the phage: Past, present and future. *Nat. Rev. Microbiol.* **13**, 777–786. https://doi.org/10.1038/nrmicro3564 (2015).
26. Lin, D. M., Koskella, B. & Lin, H. C. Phage therapy: An alternative to antibiotics in the age of multi-drug resistance. *World J. Gastrointest. Pharmacol. Ther.* **8**, 162–173. https://doi.org/10.4292/wjgpt.v8.i3.162 (2017).
27. Liu, R. Y. *et al.* Bacteriophage therapy in aquaculture: Current status and future challenges. *Folia Microbiol.* **67**, 573–590. https://doi.org/10.1007/s12223-022-00965-6 (2022).
28. Mu, A. *et al.* Assessment of the microbiome during bacteriophage therapy in combination with systemic antibiotics to treat a case of staphylococcal device infection. *Microbiome* **9**, 92. https://doi.org/10.1186/s40168-021-01026-9 (2021).
29. Febvre, H. P. *et al.* PHAGE study: Effects of supplemental bacteriophage intake on inflammation and gut microbiota in healthy adults. *Nutrients* **11**, 666. https://doi.org/10.3390/nu11030666 (2019).
30. Dissanayake, U., Ukhanova, M., Moye, Z. D., Sulakvelidze, A. & Mai, V. Bacteriophages reduce pathogenic *Escherichia coli* counts in mice without distorting gut microbiota. *Front. Microbiol.* **10**, 1984. https://doi.org/10.3389/fmicb.2019.01984 (2019).
31. Hsu, B. B. *et al.* Dynamic modulation of the gut microbiota and metabolome by bacteriophages in a mouse model. *Cell Host Microbe* **25**, 803. https://doi.org/10.1016/j.chom.2019.05.001 (2019).
32. Veyrand-Quiros, B. *et al.* Use of bacteriophage vB_Pd_PDCC-1 as biological control agent of *Photobacterium damselae* subsp. *damselae* during hatching of longfin yellowtail (*Seriola rivoliana*) eggs. *J. Appl. Microbiol.* **129**, 1497–1510. https://doi.org/10.1111/jam.14744 (2020).
33. Donati, V. L., Madsen, L., Middelboe, M., Strube, M. L. & Dalsgaard, I. The gut microbiota of healthy and *Flavobacterium psychrophilum*-infected rainbow trout fry is shaped by antibiotics and phage therapies. *Front. Microbiol.* **13**, 771296. https://doi.org/10.3389/fmicb.2022.771296 (2022).
34. Pulkkinen, K. *et al.* Intensive fish farming and the evolution of pathogen virulence: The case of columnaris disease in Finland. *Proc. R. Soc. B Biol. Sci.* **277**, 593–600. https://doi.org/10.1098/rspb.2009.1659 (2010).
35. Holt, R. A., Sanders, J. E., Zinn, J. L., Fryer, J. L. & Pilcher, K. S. Relation of water temperature to *Flexibacter columnaris* infection in steelhead trout (Salmo-Gairdneri), coho (*Oncorhynchus kisutch*) and chinook (*Oncorhynchus tshawytscha*) salmon. *J. Fish. Res. Board Can.* **32**, 1553–1559. https://doi.org/10.1139/f75-182 (1975).
36. Declercq, A. M., Haesebrouck, F., Van den Broeck, W., Bossier, P. & Decostere, A. Columnaris disease in fish: A review with emphasis on bacterium-host interactions. *Vet. Res.* **44**, 27. https://doi.org/10.1186/1297-9716-44-27 (2013).
37. Evenhuis, J. P., LaPatra, S. E. & Marancik, D. Early life stage rainbow trout (*Oncorhynchus mykiss*) mortalities due to *Flavobacterium columnare* in Idaho, USA. *Aquaculture* **418**, 126–131. https://doi.org/10.1016/j.aquaculture.2013.09.044 (2014).
38. Laanto, E., Bamford, J. K. H., Ravantti, J. J. & Sundberg, L. R. The use of phage FCL-2 as an alternative to chemotherapy against columnaris disease in aquaculture. *Front. Microbiol.* **6**, 829. https://doi.org/10.3389/fmicb.2015.00829 (2015).
39. Almeida, G. M. F. *et al.* The fate of bacteriophages in recirculating aquaculture systems (RAS)—Towards developing phage therapy for RAS. *Antibiotics (Basel)* **8**, 192. https://doi.org/10.3390/antibiotics8040192 (2019).
40. Kunttu, H. M. T., Runtuvuori-Salmela, A., Middelboe, M., Clark, J. & Sundberg, L. R. Comparison of delivery methods in phage therapy against *Flavobacterium columnare* infections in rainbow trout. *Antibiotics (Basel)* **10**, 914. https://doi.org/10.3390/antibiotics10080914 (2021).
41. Canny, S. G. D. *et al.* A novel gnotobiotic experimental system for Atlantic salmon (*Salmo salar* L.) reveals a microbial influence on mucosal barrier function and adipose tissue accumulation during the yolk sac stage. *Front. Cell. Infect. Microbiol.* **12**, 1068302. https://doi.org/10.3389/fcimb.2022.1068302 (2023).

42. Klindworth, A. *et al.* Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res.* **41**, 1. https://doi.org/10.1093/nar/gks808 (2013).
43. Nordgard, A. S. R. *et al.* Anaerobic digestion of pig manure supernatant at high ammonia concentrations characterized by high abundances of Methanosaeta and non-euryarchaeotal archaea. *Sci. Rep.* **7**, 15077. https://doi.org/10.1038/s41598-017-14527-1 (2017).
44. Edgar, R. C. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**, 2460–2461. https://doi.org/10.1093/bioinformatics/btq461 (2010).
45. Edgar, R. C. UNOISE2: Improved error-correction for Illumina 16S and ITS amplicon sequencing. *bioRxiv* 081257. https://doi.org/10.1101/081257 (2016).
46. Edgar, R. C. SINTAX: A simple non-Bayesian taxonomy classifier for 16S and ITS sequences. *bioRxiv* 074161. https://doi.org/10.1101/074161 (2016).
47. Cole, J. R. *et al.* Ribosomal Database Project: Data and tools for high throughput rRNA analysis. *Nucleic Acids Res.* **42**, D633–D642. https://doi.org/10.1093/nar/gkt1244 (2014).
48. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic Local Alignment Search Tool. *J. Mol. Biol.* **215**, 403–410. https://doi.org/10.1006/jmbi.1990.9999 (1990).
49. R: A language and environment for statistical computing (R Foundation for Statistical Computing, 2022).
50. McMurdie, P. J. & Holmes, S. phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One* **8**, e61217. https://doi.org/10.1371/journal.pone.0061217 (2013).
51. vegan: Community Ecology Package v. 2.6.2 (2023).
52. Lucas, R. *et al.* A critical evaluation of ecological indices for the comparative analysis of microbial communities based on molecular datasets. *FEMS Microbiol. Ecol.* **93**, fiw209. https://doi.org/10.1093/femsec/fiw209 (2017).
53. Hill, M. O. Diversity and evenness: A unifying notation and its consequences. *Ecology* **54**, 427–432. https://doi.org/10.2307/1934352 (1973).
54. Schliep, K. P. phangorn: Phylogenetic analysis in R. *Bioinformatics* **27**, 592–593. https://doi.org/10.1093/bioinformatics/btq706 (2011).
55. Anderson, M. J. A new method for non-parametric multivariate analysis of variance. *Austral Ecol.* **26**, 32–46. https://doi.org/10.1111/j.1442-9993.2001.01070.pp.x (2001).
56. Murray, C. J. L. *et al.* Global burden of bacterial antimicrobial resistance in 2019: A systematic analysis. *Lancet* **399**, 629–655. https://doi.org/10.1016/S0140-6736(21)02724-0 (2022).
57. Furfaro, L. L., Payne, M. S. & Chang, B. J. Bacteriophage therapy: Clinical trials and regulatory hurdles. *Front. Cell Infect. Microbiol.* **8**, 376. https://doi.org/10.3389/fcimb.2018.00376 (2018).
58. Ganeshan, S. D. & Hosseinidoust, Z. Phage therapy with a focus on the human microbiota. *Antibiotics (Basel)* **8**, 131. https://doi.org/10.3390/antibiotics8030131 (2019).
59. Silva, Y. J. *et al.* Biological control of *Aeromonas salmonicida* infection in juvenile Senegalese sole (*Solea senegalensis*) with Phage AS-A. *Aquaculture* **450**, 225–233. https://doi.org/10.1016/j.aquaculture.2015.07.025 (2016).
60. Kim, J. H. *et al.* Biological control of *Aeromonas salmonicida* subsp. *salmonicida* infection in rainbow trout (*Oncorhynchus mykiss*) using *Aeromonas* phage PAS-1. *Transbound. Emerg. Dis.* **62**, 81–86. https://doi.org/10.1111/tbed.12088 (2015).
61. Feng, C. *et al.* Lytic bacteriophage PZL-Ah152 as biocontrol measures against lethal *Aeromonas hydrophila* without distorting gut microbiota. *Front. Microbiol.* **13**, 898961. https://doi.org/10.3389/fmicb.2022.898961 (2022).
62. Verner-Jeffreys, D. W. *et al.* Furunculosis in Atlantic salmon (*Salmo salar* L.) is not readily controllable by bacteriophage therapy. *Aquaculture* **270**, 475–484. https://doi.org/10.1016/j.aquaculture.2007.05.023 (2007).
63. Bugten, A. V. *et al.* Changes in rearing water microbiomes in RAS induced by membrane filtration alters the hindgut microbiomes of Atlantic salmon (*Salmo salar*) parr. *Aquaculture* **548**, 737661. https://doi.org/10.1016/j.aquaculture.2021.737661 (2022).
64. Vestrum, R. I., Attramadal, K. J. K., Vadstein, O., Gundersen, M. S. & Bakke, I. Bacterial community assembly in Atlantic cod larvae (*Gadus morhua*): Contributions of ecological processes and metacommunity structure. *FEMS Microbiol. Ecol.* **96**, fiaa163. https://doi.org/10.1093/femsec/fiaa163 (2020).
65. Vestrum, R. I. *et al.* Rearing water treatment induces microbial selection influencing the microbiota and pathogen associated transcripts of cod (*Gadus morhua*) larvae. *Front. Microbiol.* **9**, 851. https://doi.org/10.3389/fmicb.2018.00851 (2018).
66. Bakke, I., Coward, E., Andersen, T. & Vadstein, O. Selection in the host structures the microbiota associated with developing cod larvae (*Gadus morhua*). *Environ. Microbiol.* **17**, 3914–3924. https://doi.org/10.1111/1462-2920.12888 (2015).
67. Zaburlin, D., Quiberoni, A. & Mercanti, D. Changes in environmental conditions modify infection kinetics of dairy phages. *Food Environ. Virol.* **9**, 270–276. https://doi.org/10.1007/s12560-017-9296-2 (2017).
68. Kadavy, D. R. *et al.* Influence of infected cell growth state on bacteriophage reactivation levels. *Appl. Environ. Microbiol.* **66**, 5206–5212. https://doi.org/10.1128/AEM.66.12.5206-5212.2000 (2000).
69. Loc-Carrillo, C. & Abedon, S. T. Pros and cons of phage therapy. *Bacteriophage* **1**, 111–114. https://doi.org/10.4161/bact.1.2.14590 (2011).
70. Huang, J. Q. *et al.* Salmonella phage CKT1 significantly relieves the body weight loss of chicks by normalizing the abnormal intestinal microbiome caused by hypervirulent *Salmonella pullorum*. *Poultry Sci.* **101**, 101668. https://doi.org/10.1016/j.psj.2021.101668 (2022).
71. Pereira, C. *et al.* Bacteriophages with potential for inactivation of fish pathogenic bacteria: Survival, host specificity and effect on bacterial community structure. *Mar. Drugs* **9**, 2236–2255. https://doi.org/10.3390/md9112236 (2011).
72. Lamy-Besnier, Q. *et al.* Prophylactic administration of a bacteriophage cocktail is safe and effective in reducing *Salmonella enterica* serovar *typhimurium* burden in vivo. *Microbiol. Spectr.* **9**, e0049721. https://doi.org/10.1128/Spectrum.00497-21 (2021).
73. Ahasan, M. S. *et al.* Bacteriophage versus antibiotic therapy on gut bacterial communities of juvenile green turtle, *Chelonia mydas*. *Environ. Microbiol.* **21**, 2871–2885. https://doi.org/10.1111/1462-2920.14644 (2019).
74. Jakobsen, R. R. *et al.* A bacteriophage cocktail significantly reduces *Listeria monocytogenes* without deleterious impact on the commensal gut microbiota under simulated gastrointestinal conditions. *Viruses (Basel)* **14**, 190. https://doi.org/10.3390/v14020190 (2022).
75. Richards, P. J., Connerton, P. L. & Connerton, I. F. Phage biocontrol of *Campylobacter jejuni* in chickens does not produce collateral effects on the gut microbiota. *Front. Microbiol.* **10**, 476. https://doi.org/10.3389/fmicb.2019.00476 (2019).
76. Mai, V., Ukhanova, M., Reinhard, M. K., Li, M. & Sulakvelidze, A. Bacteriophage administration significantly reduces *Shigella* colonization and shedding by *Shigella*-challenged mice without deleterious side effects and distortions in the gut microbiota. *Bacteriophage* **5**, e1088124. https://doi.org/10.1080/21597081.2015.1088124 (2015).
77. Almeida, A. R., Alves, M., Domingues, I. & Henriques, I. The impact of antibiotic exposure in water and zebrafish gut microbiomes: A 16S rRNA gene-based metagenomic analysis. *Ecotoxicol. Environ. Saf.* **186**, 109771. https://doi.org/10.1016/j.ecoenv.2019.109771 (2019).
78. Uyaguari, M., Key, P., Moore, J., Jackson, K. & Scott, G. Acute effects of the antibiotic oxytetracycline on the bacterial community of the grass shrimp, *Palaemonetes pugio*. *Environ. Toxicol. Chem.* **28**, 2715–2724. https://doi.org/10.1897/08-514.1 (2009).
79. Navarrete, P., Mardones, P., Opazo, R., Espejo, R. & Romero, J. Oxytetracycline treatment reduces bacterial diversity of intestinal microbiota of Atlantic salmon. *J. Aquat. Anim. Health* **20**, 177–183. https://doi.org/10.1577/H07-043.1 (2008).

80. Willing, B. P., Russell, S. L. & Finlay, B. B. Shifting the balance: Antibiotic effects on host-microbiota mutualism. *Nat. Rev. Microbiol.* **9**, 233–243. https://doi.org/10.1038/nrmicro2536 (2011).
81. Antonopoulos, D. A. *et al.* Reproducible community dynamics of the gastrointestinal microbiota following antibiotic perturbation. *Infect. Immun.* **77**, 2367–2375. https://doi.org/10.1128/Iai.01520-08 (2009).
82. Narrowe, A. B. *et al.* Perturbation and restoration of the fathead minnow gut microbiome after low-level triclosan exposure. *Microbiome* **3**, 6. https://doi.org/10.1186/s40168-015-0069-6 (2015).
83. Almeida, A. R., Domingues, I. & Henriques, I. Zebrafish and water microbiome recovery after oxytetracycline exposure. *Environ. Pollut.* **272**, 116371. https://doi.org/10.1016/j.envpol.2020.116371 (2021).
84. Payne, C. J., Turnbull, J. F., MacKenzie, S. & Crumlish, M. Investigating the effect of an oxytetracycline treatment on the gut microbiome and antimicrobial resistance gene dynamics in Nile tilapia (*Oreochromis niloticus*). *Antibiotics (Basel)* **10**, 1213. https://doi.org/10.3390/antibiotics10101213 (2021).

Acknowledgements
We would like to thank Amalie Mathisen for valuable help in the lab. Further, we thank Rolf Myklebust from Aquagen AS for providing us with salmon eggs. We also thank Lotta-Riina Sundberg for providing us with phage FCL-2 and helpful tips on phage cultivation, as well as Jean-Marc Ghigo and David Perez-Pascual for providing us with *F. columnare* strain Fc7.

Author contributions
A.F., M.G., E.A., O.V. and I.B. designed and conceived the experiment. A.F., M.G. and T.V. conducted the experiments and A.F. and M.G. analysed the data. A.F. wrote a first draft of the manuscript, and all authors improved and approved it.

Funding
Open access funding provided by Norwegian University of Science and Technology. MSG and AWF were supported by a PhD scholarship from NTNU, Faculty of Natural Science.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-023-44987-7.

Correspondence and requests for materials should be addressed to I.B.

Reprints and permissions information is available at www.nature.com/reprints.

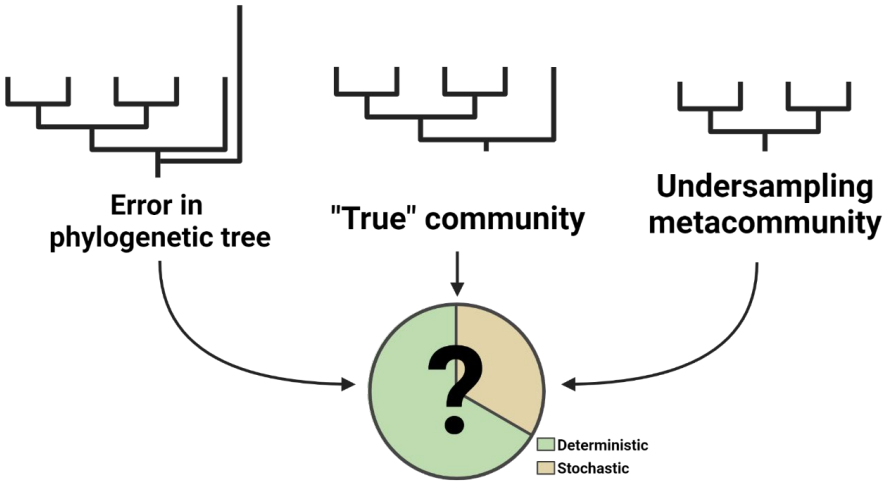
Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Work I

An exploration of the impact of phylogenetic tree structure
on NTI and βNTI estimates of community assembly

An exploration of the impact of phylogenetic tree structure on NTI and βNTI estimates of community assembly

Madeleine J. S. Gundersen, 2023-07-05



An exploration of the impact of phylogenetic tree structure on NTI and βNTI estimates of community assembly.

Outline

Background

Quantifying community assembly is critical to understanding how patterns of community diversity arise. Bacterial communities in a local community are thought to be structured by the four higher-order community assembly processes of selection, drift, dispersal and diversification (Nemergut et al., 2013; Vellend, 2010). Selection is a deterministic process based on fitness differences between populations in response to environmental conditions (i.e. environmental filtering) and biotic interactions. In contrast, drift is a stochastic process that occurs due to random mortality events that manifest as unpredictable fluctuations in population abundance. Furthermore, through the process of dispersal, bacteria migrate from the regional or metacommunity to the local community and, over evolutionary time, will undergo diversification, creating new genetic lineages (Nemergut et al., 2013; Vellend, 2010, Figure 1).

A key challenge in quantifying community assembly is the statistical quantification of stochastic community assembly, as such processes by definition result in the absence of patterns in the dataset (Zhou & Ning, 2017). Null model-based approaches have gained popularity, as stochastic community assembly patterns can be generated from the properties of the dataset under study. Two popular null models are the Nearest Taxon Index (NTI) and βNTI (Beta Nearest Taxon Index). Briefly, the NTI measures the degree of phylogenetic clustering or overdispersion in a single community (i.e. within a sample). Highly clustered samples are assumed to be deterministically structured, e.g. by environmental filtering, with the underlying assumption that phylogenetically close populations thrive in the same environments (Zhou & Ning, 2017). βNTI measures the phylogenetic similarity between two communities (i.e. two samples). If two communities are more or less phylogenetically similar than expected by chance, it is assumed that the communities have been structured by selection. Thus, NTI and βNTI are metrics that rely on a phylogenetic tree to assess co-occurrence patterns and elucidate underlying community assembly processes (Stegen et al., 2013).

NTI measures the degree of phylogenetic clustering in a single community (Webb et al., 2002). The degree of clustering depends on the structure of the phylogenetic tree estimated from the

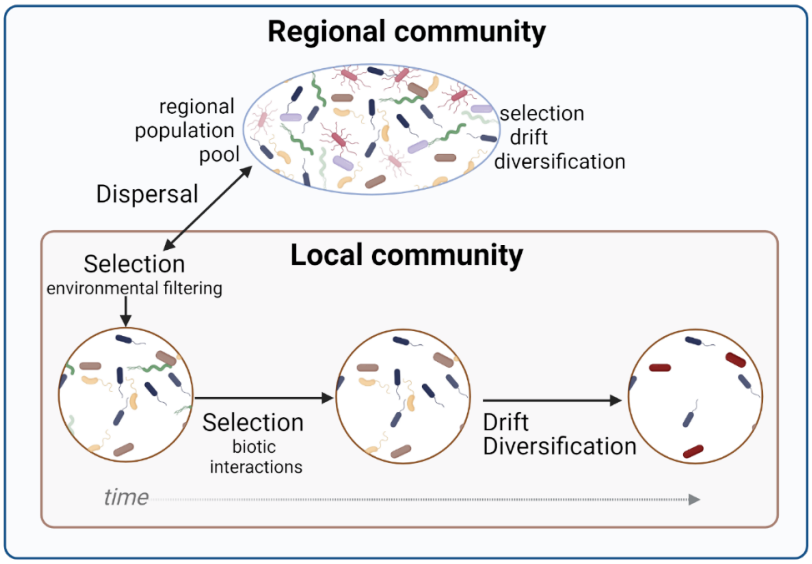


Figure 1: The processes of selection, drift, dispersal and diversification shape regional and local communities over time. Dispersal adds individuals from the regional population pool to the local community. Population establishment is determined by selection through environmental filtering. Over time, community composition within the local community is shaped by selection (environmental filtering and biotic interactions), drift and diversification. Diversification (i.e. speciation) can create new populations, while selection and drift can reduce the abundance of populations or cause them to go extinct. Figure inspired from Vellend, 2010 and 2016.

regional community. To estimate NTI, multiple null communities are generated using a null model approach. To generate null communities, community properties such as richness and evenness are kept stable, but population labels are randomised in the phylogenetic tree. This randomisation is assumed to generate stochastically assembled communities. Typically, 1000 null communities are generated. NTI is based on comparing the observed phylogenetic distance between co-occurring populations in a community with the average phylogenetic distance of the null communities. During estimation, the pairwise phylogenetic distance between populations in the community ($MNTD_{observed}$) and the mean phylogenetic distances for the null communities (mean $MNTD_{null}$) are calculated. The NTI is then calculated as the difference between the mean $MNTD_{null}$ and $MNTD_{observed}$ divided by the standard deviation of the null distribution of $MNTD_{null}$.

NTI values greater than zero indicate phylogenetic clustering (i.e. species are more closely related than expected by chance). Conversely, values less than zero indicate phylogenetic overdispersion (i.e. species are more distantly related than expected by chance). βNTI is calculated similarly to NTI but estimates the average phylogenetic distance between two communities rather than within a community (Stegen et al., 2012). For NTI and βNTI, two standard deviations from the null model are usually considered significant. As the null models rely heavily on the phylogenetic tree of the metacommunity, there are several pitfalls that can occur with these approaches (Figure 2).

In this work, I investigated the effect of 1) an incorrect phylogenetic tree and 2) the size of the metacommunity on the results of the NTI and βNTI analyses. Datasets based on 16S rRNA gene amplicon sequencing may contain sequences from organisms other than bacteria, such as archaea and eukaryotic ribosomal sequences. Such erroneous sequences may be classified as bacteria during the bioinformatic processing of the sequencing data. Many bacterial ecologists rely on the bioinformatic pipelines to identify the erroneous sequences and filter the dataset to include populations classified as bacteria. Generally, unless the population appears suspicious, its taxonomic classification is not confirmed with additional bioinformatics (e.g. BLAST or RDP classifiers). The presence of such erroneous sequences in the dataset will ultimately result in a phylogenetic tree containing outliers and is expected to have a major impact on the null model estimates. To my knowledge, the extent to which this affects estimates of the community assembly process has not been investigated.

Furthermore, the null models NTI and βNTI depend on the selected metacommunity as it affects the size and structure of the phylogenetic tree. However, there are no guidelines to define the cut-off for the regional species pool. It is well documented that below a cutoff, increased sampling effort increases the regional species pool (Nemergut et al., 2013). Thus, there is a high probability that mean phylogenetic relationships will change as a function of sampling effort. Few samples are expected to represent a smaller proportion of the regional species pool. The consequence of low sampling effort is that the null model communities will, on average, have a higher similarity to the observed community, making it less likely to obtain significantly different

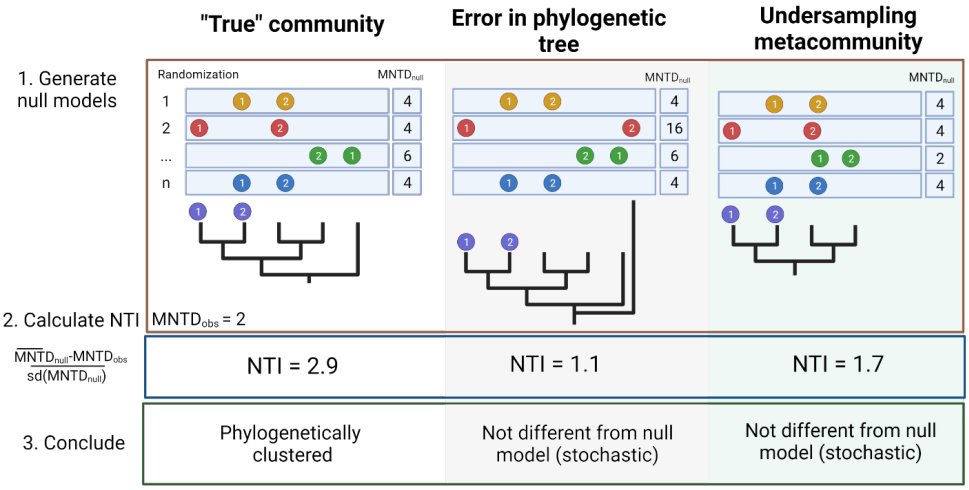


Figure 2: Pitfalls when calculating the nearest taxon unit (NTI). In this simplified example, a community contains two closely related populations (1 and 2), separated on the phylogenetic tree by two branches each with an arbitrary length of one, resulting in a mean phylogenetic distance ($MNTD_{observed}$) of 2 (1 + 1). Null modelling is performed by randomly shuffling the populations on the tips of the phylogenetic tree, and the $MNTD_{null}$ is calculated for each randomization. The estimated NTI value indicates that the community is more phylogenetically clustered than expected by chance (NTI = 2.9 > 2). However, if a sequence not representing the bacterial community is included in the phylogenetic tree, some randomisations will be erroneous and substantially increase the standard deviation of the null models (e.g. randomisation 2 with red dots). This error results in that the community no longer is phylogenetically clustered (NTI = 1.1 < 2). Similarly, undersampling the regional species pool truncates the phylogenetic tree, and the null model is too similar to the observed community (NTI = 1.7 < 2).

$MNTD_{observed}$ or $\beta MNTD_{observed}$. Defining the boundaries of the metacommunity is critical. For example, more bacterial taxa are expected to be observed from a range of spatial locations than from just one geographical location. If all samples are included in the regional species pool, the null models may be too different from the actual samples of interest, leading to a bias towards significant differences between the observed and null model communities. Therefore, it is valuable to explore and understand the impact of metacommunity choice on community assembly analysis tools.

Datasets used to explore pitfalls in phylogenetically based analysis

Two datasets were used to explore the pitfalls related to the estimation of the assembly processes. Both datasets investigated fish-microbiome interactions and sampled fish guts, rearing- and intake water (Mathisen, 2019; Vestrum et al., 2020).

Dataset 1: Salmon alevins reared with *r*- or *K*-selected intake water

In Dataset 1, Atlantic salmon eggs were disinfected and hatched germ-free in sterile medium in ventilated-cell culture flasks (Mathisen, 2019). The rearing water was changed three times a week with sterile medium during the first week and with lake water thereafter. Two types of lake water were used. Half of the flasks received lake water as collected (*K*-selected) and half received nutrient-enriched lake water (*r*-selected). After 22 days (15 days exposure to bacteria), the bacterial communities of the water (1 sample per flask) and larval guts (4 individuals per flask) were analysed using 16S rRNA gene amplicon sequencing. In addition, the intake water was sampled on days 14, 16, 18 and 21.

An ASV table was generated from the sequencing reads using the USEARCH pipeline and the RDP v16 database. The dataset contained 88 samples (8 intake water, 16 rearing water and 64 gut samples) and 3227 ASVs. The ASV table was manually quality assessed by the investigators to exclude contaminants based on library preparation controls and ASVs likely to represent salmon genes. After removal of contamination, the ASV table was scaled to 11 000 reads per sample and rounded. This quality-assessed and normalised dataset is referred to as Dataset 1 and was used as the input for the following exploratory analysis.

In Dataset 1, eucaryotic sequences had been classified as bacteria by the SINTAX command in USEARCH. These errors were not detectable in the taxonomy table as all ASVs were classified to the domain Bacteria. However, when inspecting the phylogenetic tree, I identified several ASVs with sequences that diverged from most other ASV sequences (Figure 3a). I performed a BLAST search on the ASVs that diverged from the other sequences and found that most of these had high similarity to salmonid fish (Supplementary Table 1). When these 37 ASVs were removed, the

phylogenetic tree appeared to only display bacteria, as the tree was uniform with no outliers Figure 3b). An R script to identify outlier sequences is provided in the supplementary material.

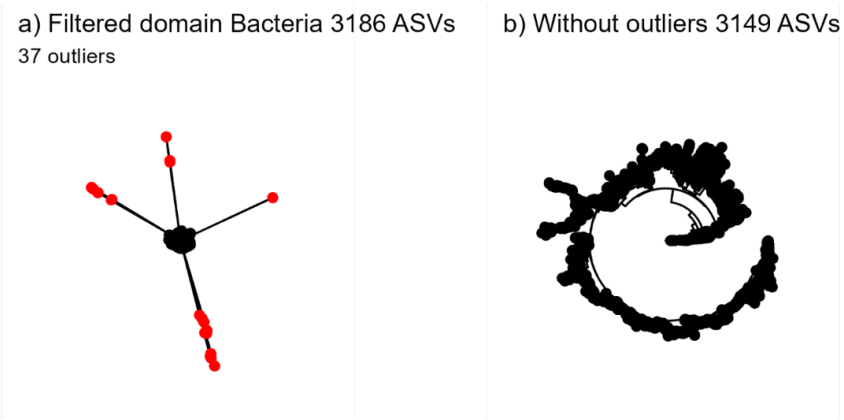


Figure 3: Phylogenetic trees of Dataset 1 (Salmon alevins) filtered to contain a) only ASVs belonging to the domain bacteria (3186 ASVs) and b) without erroneous sequences (3149 ASVs).

Dataset 2: Cod larvae reared in different water treatment systems

In Dataset 2, Atlantic cod eggs were disinfected and divided into nine rearing tanks. Triplicate rearing tanks received water from one of three different water treatment systems. The water treatment systems were a flow-through system (FTS), a microbially matured water system (MMS) or a recirculating water system (RAS) (Vestrum et al., 2020). The eggs hatched in the rearing tanks and received water from the different systems for 30 days. On day 30, all tanks were switched over to receive water from the MMS system. The experiment was terminated after 46 days.

Throughout the experiment, bacterial community samples of the intake- and rearing water, feed and cod larval gut were taken. 16S rRNA gene amplicon sequencing was used to analyse the communities. The sequencing reads were processed similarly to Dataset 1, but an OTU table was generated instead of an ASV table. The dataset contained 197 samples (108 cod gut-, 48 rearing water, 13 intake water and 28 feed samples) and 3336 OTUs. OTUs representing algae, Archaea, *Cyanobacteria*, *Chloroplasts* and the contaminant *Propionibacterium acne* from the OTU table

before each sample was scaled to 12 100 sequencing reads. This normalised OTU table represents Dataset 2 which was used for as input in the following explorative analysis.

As with Dataset 1, some OTUs were misclassified in the bioinformatical pipelines. These 42 outliers were examined through BLAST and were found to have high similarity to Atlantic cod and several different types of fungi (Figure 4, Supplementary Table 2). These outliers were subsequently removed.

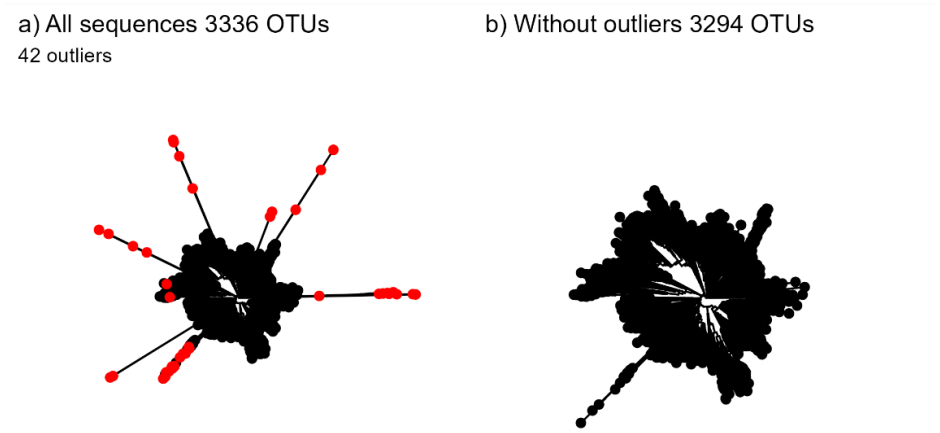


Figure 4: Phylogenetic trees of Dataset 2 (Cod larvae) with a) all OTUs (3336 OTUs) and b) without outlier sequences (3294 OTUs).

The presence of a phylogenetic signal in both datasets without the outliers was confirmed as required before estimating NTI and βNTI (Stegen et al., 2013, Supplementary Figure 1). To test for a phylogenetic signal I first calculated the abundance-weighted mean for environmental variables, which was standardised to the maximum phylogenetic distance. Then, a multivariate Mantel correlogram was calculated based on the standardised weighted mean and phylogenetic distances using the function *mantel.correlog()* from the R-package vegan (version 2.6-4, Oksanen et al., 2022). A Mantel statistical test was used to infer statistical significance, which was Holm-corrected with 1000 permutations.

Errors in the phylogenetic tree

Effect of outlier sequences on community characteristics

In general, removal of the erroneous sequences affected community metrics that use the phylogenetic tree as input. The 37 outliers in Dataset 1 and the 42 outliers in Dataset 2 were a marginal fraction of the population pool and were relatively rare in each sample (Table 1).

Table 1: Community characteristics in Dataset 1 and 2. The outlier sequences identified did not make up much of the population pool. The outliers did not contribute significantly to the species pool on a sample basis and generally contributed little to the overall relative abundance in each sample.

	n outliers	% of population pool	Max ASV sample loss	Relative abundance per sample (± SD)	Max sample abundance
Dataset 1	37	1.2%	8	Water: 0.1% (±0.06%) Gut: 0.7% (±1.5%)	9.0%
Dataset 2	42	1.3%	5	Water&feed: 0.04% (±0.1%) Gut:0.06% (±0.24%)	2.4%

To test how removing the outlier sequences affected the community composition in each sample, I quantified the similarity between the full dataset and the one without outlier sequences. Similarity was quantified for each sample by comparing the sample community composition in the full dataset with the composition after removing the outliers. As similarity indices I used Bray-Curtis, Sørensen, unweighted (uw) UniFrac and weighted (w) UniFrac. Bray-Curtis quantifies the similarity in community composition and takes into account differences in relative abundance. The Sørensen similarity quantifies the similarity in the presence-absence of populations (Legendre & Cáceres, 2013). The UniFrac indices incorporate both the composition of the bacterial community and the phylogenetic relationship between populations (Lozupone & Knight, 2005). The unweighted UniFrac does not incorporate relative abundance, while the weighted UniFrac does. The similarities were calculated using *distance*() from phyloseq (version 1.42.0, (McMurdie & Holmes, 2013).

I found that Bray Curtis, Sørensen and weighted UniFrac were largely unaffected by removing the outliers (Table 2, Figure 5). However, the effect of removing the erroneous sequences was drastic for the unweighted UniFrac similarity with 93.2% and 25.9% of the samples in Dataset 1 and 2, respectively, having a similarity below 0.95. Further, there was a trend that similarity increased

more when the sample contained more outlier populations. This reduction clearly illustrates the enormous effect errors in the phylogenetic tree can have.

Thus, investigations of the community properties in Dataset 1 and 2 indicate that non-phylogenetic tree based analytical approaches are unaffected by the presence of some outlier sequences. However, the outliers introduce bias when the phylogenetic tree is crucial to the analytical framework, and especially when only the presence-absence of populations is evaluated.

Table 2: The average±SD similarity using different similarity indices comparing each sample composition between the full dataset and the one without outliers.

	BC	Sørensen	u UniFrac	w UniFrac	n uw UniFrac < 0.95
Dataset 1	1.00±0.01	0.99±0.03	0.37±0.22	0.97±0.07	82 of 88
Dataset 2	1.00±0.001	0.99±0.004	0.37±0.22	0.97±0.07	51 of 197

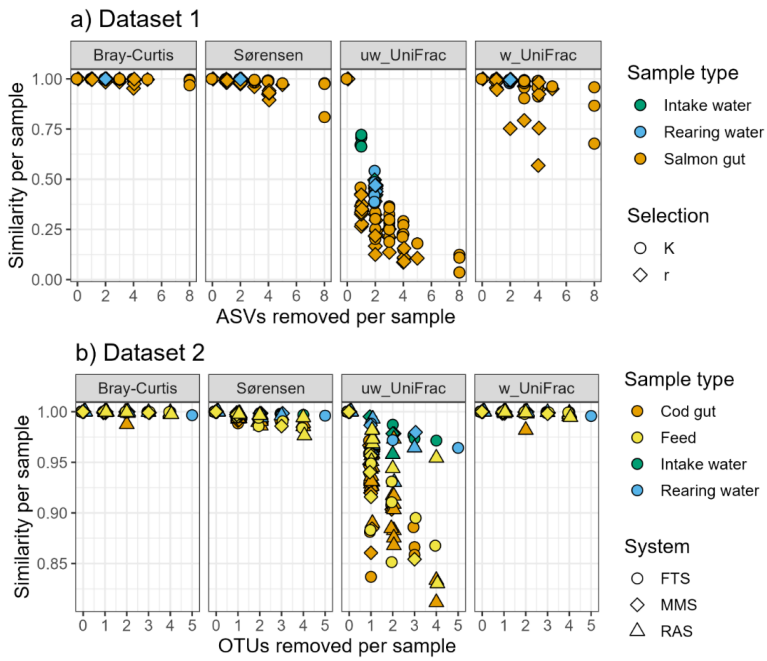


Figure 5: The Bray-Curtis, Sørensen, unweighted UniFrac and weighted UniFrac similarity for each sample between the datasets with and without the outliers in a) Dataset 1 and b) Dataset 2.

Effect of outliers on NTI

NTI was estimated for each gut, intake and rearing water sample in the datasets with and without the outlier sequences. In general, removal of the outliers increased the NTI (Figure 6).

In both datasets, the NTI had increased in 98.8% of the samples. On average, NTI had increased by 6.1±2.8, 8.1±2.2 and 12.4±1.1 in the gut, intake and rearing water samples, respectively, in Dataset 1. For the water samples, the conclusion that the communities were phylogenetically clustered remained the same, although the strength of the conclusion was stronger when the erroneous sequences were removed (**Table 3**). In total, 82.8% of the samples had a change in conclusion, and 9.4% of the gut samples had a drastic change in conclusion from phylogenetically overdispersed to clustered. In Dataset 2, the NTI did not change as drastically. On average, the NTI increased by 1.3±0.65 and all samples were phylogenetically clustered in both the full and outlier-free datasets. However, there was a clear tendency that removing the outliers consistently increased the NTI and therefore underestimated the NTI. Thus, if the samples had been closer to the null model, more samples might have changed from stochastic to clustered.

Overall, errors in the phylogenetic tree can lead to substantially different interpretations of the underlying ecological mechanisms.

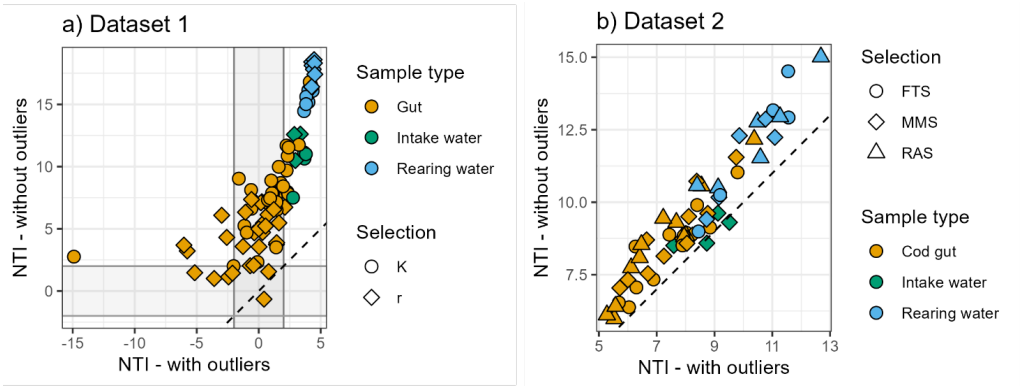


Figure 6: Differences in estimated NTI for each sample in the data with and without outliers in a) data set 1 (salmon alevin) and b) data set 2 (cod larvae). The dashed line shows the 1:1 relationship. Values within the grey area are within two standard deviations of the null model.

Table 3: The NTI was calculated in the datasets with and without outlier sequences. For Dataset 1, the NTI was estimated within each sample type and inlet water selection regime (r- and K-selection). For Dataset 2, the NTI was estimated within sample types and water treatment systems (FTS, MMS and RAS). For each sample type, the βNTI conclusion was assessed and the number of times (n) the analytical conclusion changed was counted (orange). For each sample type, the percentage of n is given (% comparisons). D = phylogenetically dispersed, C = phylogenetically clustered, S = stochastic.

	Sample Type	% changed	Conclusion Full	Conclusion wo/ Outliers	n	% of sample comparisons
Dataset 1	Salmon gut	82.8	D	C	6	9.3
			D	S	4	6.3
			S	C	43	67.2
			C	C	9	14.1
			S	S	2	3.1
	Rearing water	0	C	C	16	100
	Intake water	0	C	C	8	100
Dataset 2	Cod gut	0	C	C	108	100
	Rearing water	0	C	C	13	100
	Intake water	0	C	C	48	100

Effect of outliers on βNTI

The βNTI was estimated for gut, intake and rearing water in Dataset 1 and for gut and rearing water in Dataset 2. In Dataset 1, most of the estimated βNTI changed when the outlier sequences were removed from the dataset, but in Dataset 2 the βNTI did not change significantly (Table 4, Figure 7). For the water samples, the general trend was that βNTI decreased when the outliers were removed. This resulted in 66.1% of the rearing water and 25.0% of the intake water sample comparisons in Dataset 1 and 6.7% of the rearing water sample comparisons in Dataset 2 changing the conclusion from stochastic to homogeneous selection in Dataset 1 (Table 4).

For the gut samples, the general trend was that the βNTI increased in Dataset 1 but decreased in Dataset 2. In Dataset 1, 91.2% of the gut samples had higher βNTI. These increases caused the conclusion to change from stochastic to heterogeneous selection dominating the community assembly in 43.3% of the comparisons. In Dataset 2, 78.7% of the cod gut sample comparisons decreased in βNTI and 6.3% of the sample comparisons changed the conclusion.

Table 4: The βNTI was calculated in the datasets with and without outlier sequences. For Dataset 1, the βNTI was estimated within each sample type (gut, intake and rearing water) and intake water selection regime (r and K selection). For dataset 2, βNTI was estimated within each sample type (gut and rearing water) and water treatment regime (FTS, MMS and RAS). For each sample type, the βNTI conclusion was assessed and the number of times (n) the analysis conclusion changed was counted (orange). For each sample type, the percentage of n is given (% comparisons). HeS = heterogeneous selection, HoS = homogeneous selection, S = stochastic.

	Sample Type	% changed	Conclusion full	Conclusion wo/ outliers	n	% comparisons
Dataset 1	Salmon gut	44.2	HeS	S	9	0.9
			S	HeS	430	43.3
			HeS	HeS	13	1.3
			S	S	540	54.4
	Intake water	25.0	S	HoS	3	25.0
			HoS	HoS	1	8.3
			S	S	8	66.7
	Rearing water	66.1	S	HoS	37	66.1
			HoS	HoS	18	32.1
			S	S	1	1.8
Dataset 2	Cod gut	6.3	HoS	S	2	0.5
			S	HeS	2	0.5
			S	HoS	23	5.3
			HoS	HoS	220	50.9
			S	S	185	42.8
	Rearing water	6.7	S	HoS	3	6.7
			HoS	HoS	32	71.1
			S	S	10	22.2

In Dataset 1, the effect of erroneous sequences was drastic, with 45.2% of all sample comparisons leading to erroneous conclusions. In Dataset 2, the effect was less pronounced, with 6.3% of the sample comparisons changing the conclusion. Nevertheless, these results indicate that errors in the phylogenetic tree can have a major impact on the interpretation of the data.

Choice of metacommunity

As the regional population pool affects the phylogenetic tree, I investigated the effect of metacommunity size on NTI and βNTI. In Dataset 1 (salmon alevin) there were three different sample types; salmon gut, intake and rearing water. Dataset 2 (cod larvae) had four different sample types: cod gut, intake and rearing water and feed. It is not unreasonable to imagine that

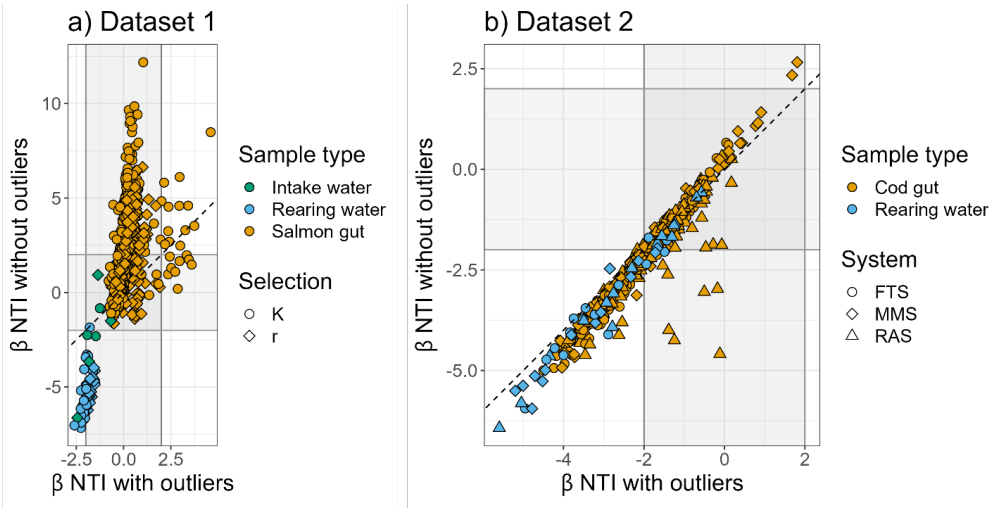


Figure 7: Differences in estimated βNTI between pairs of samples based on data with and without outliers in a) Dataset 1 (salmon alevin) and b) Dataset 2 (cod larvae). The dashed line shows the 1:1 relationship. Values within the grey area are within two standard deviations of the null model. Only sample comparisons within the same sample type and sampling water selection regime are shown.

the experiments could have been carried out without sampling the intake or rearing water or the feed, as the main focus of these experiments was on the bacterial communities of the fish. It is expected that more sample types will result in a larger regional community. Therefore, to explore the effect of metacommunity size, I excluded some sample types and investigated how a smaller regional population pool affected the NTI and βNTI estimates.

For both datasets 1 and 2, I used the dataset with no outliers as the initial metacommunity. I then filtered out sample types to obtain a metacommunity consisting of rearing water and gut samples only (rearing+gut metacommunity) or gut samples only (gut metacommunity). In Dataset 1, removal of rearing water and all water samples from the metacommunity resulted in a loss of 20.3% (640 ASVs) and 33.1% (1043 ASVs) of ASVs, respectively (Table 5). In Dataset 2, removal of **f**eed and intake water resulted in a loss of 7.0% (230 OTUs) and 42.8% (1411 OTUs) of the regional OTU pool, respectively (Table 5).

Table 5: Overview of how the metacommunity size affected the number of ASVs or OTUs in the regional population pool and the average estimated NTI and βNTI (average± standard deviation) for the gut and rearing water samples.

	Metacommunity	All samples	Rearing+Gut	Gut
Dataset 1	n ASVs	3149	2509	2106
	NTI gut samples	6.07 ±3.11	5.16±2.84	5.49±2.97
	NTI rearing water	16.06±1.34	14.27±0.86	-
	βNTI gut samples	2.10±1.93	2.55±2.02	2.41±1.99
	βNTI rearing samples	-5.37±1.08	-4.41±1.05	-
Dataset 2	n OTUs	3294	3064	1883
	NTI gut samples	8.43±1.70	8.15±1.63	7.13±1.31
	NTI rearing water	11.60±1.66	11.08±1.54	-
	βNTI gut samples	-2.24±1.28	-2.15±1.28	-1.61±1.21
	βNTI rearing samples	-3.45±1.52	-3.31±1.46	-

Effect of metacommunity size on NTI

In general, reducing the size of the metacommunity resulted in lower NTI (Figure 8a and c). In Dataset 1, the rearing water had an average of 2.34±0.66 lower NTI than when all samples were included in the metacommunity (Figure 8a). Although the NTI was reduced, the conclusion that the rearing water was phylogenetically clustered remained (Figure 8b). The reduction in NTI was less pronounced in the salmon gut samples, with an average NTI reduction of -0.91±0.35 and -0.58±0.26 when the intake and all water samples were removed from the metacommunity, respectively. Only for 6.3% of the gut samples the conclusion changed. These samples went from being phylogenetically clustered when all ASVs were present in the metacommunity to being stochastic when the metacommunity was reduced.

When the feed and intake water samples were removed from the metacommunity in Dataset 2, NTI decreased on average by -0.51±0.40 in the water samples and -0.28±0.27 in the gut samples (Figure 8c). Not surprisingly, the decrease was higher with an average of -1.30±0.50 when only the gut samples were present. Changing the metacommunity size did not change the conclusion based on the NTI values, as all samples remained phylogenetically clustered (Figure 8d).

In conclusion, the reduced metacommunity size resulted in the communities being evaluated as less phylogenetically structured.

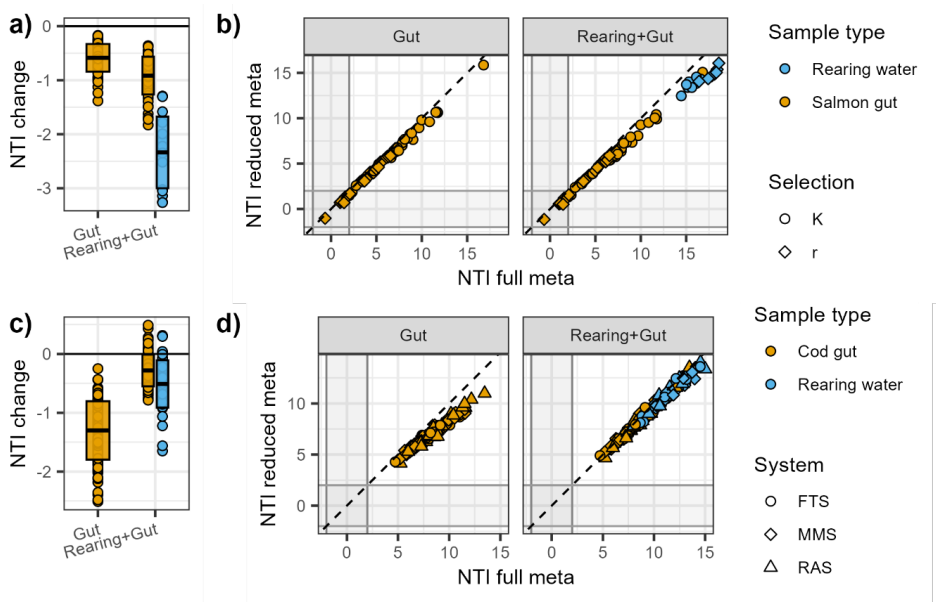


Figure 8: Estimated NTI values changed when the metacommunity was reduced. The change in estimated NTI for all gut and rearing water samples was calculated as the difference in estimated NTI between the reduced and the original metacommunity in a) Dataset 1 and c) Dataset 2. The metacommunity is indicated on the x-axis. The box plot shows the mean change in NTI±standard deviation. Comparison of NTI in the reduced and full metacommunity in b) Dataset 1 and d) Dataset 2. The dashed line indicates a 1:1 relationship. Values within the grey areas were within two standard deviations of the null model.

Effect of metacommunity size on βNTI

In general, a smaller metacommunity resulted in increased βNTI estimates (Figure 9). The rearing water βNTI increased on average by 0.96±0.19 in Dataset 1 and 0.14±0.15 in Dataset 2 when only the rearing water and gut samples were part of the metacommunity (Figure 9a and c). These changes resulted in only one sample comparison in Dataset 2 changing the conclusion from homogeneous selection to stochastic. Otherwise, all other sample comparisons had the same ecological interpretation.

In Dataset 1, the salmon gut sample comparisons increased in βNTI by an average of 0.45±0.17 when the metacommunity consisted of the rearing water and gut samples, and by 0.31+0.16 when it contained only the gut samples. This resulted in 9.2% and 6.4% of the gut sample

comparisons changing the conclusion from stochastic to heterogeneous selection. In data set 2, the cod gut samples also increased the βNTI. When only rearing water and gut samples were part of the metacommunity, βNTI increased by 0.090±0.097, but increased by 0.64±0.22 when only the gut was part of the metacommunity. These increases resulted in 2.1% and 17.1% of the sample comparisons changing from homogeneous selection to stochastic community assembly, respectively.

In conclusion, the estimated βNTI increased when parts of the metacommunity were removed, and thus the size of the metacommunity affected the conclusions drawn from the βNTI framework in both datasets examined. The observations suggest that a more diverse metacommunity increases the likelihood of an ecological process being categorised as deterministic, and similarly, undersampling increases the likelihood of a comparison being categorised as stochastic. Nevertheless, the changes in βNTI were relatively small, with a maximum difference of about 1.5. Thus, in terms of interpreting ecological processes, the sample comparisons close to the null model are most affected. Therefore, I recommend testing the robustness of the conclusions drawn by filtering out parts of the metacommunity under study and reanalysing the dataset.

Conclusion and recommendations

Overall, this exploratory data analysis showed that changes in the phylogenetic tree could impact the estimated NTI and βNTI. I observed that errors in the phylogenetic three had the most impact on the null model analysis, and had a stronger effect size on changing the interpretation of the ecological processes than changes in metacommunity structure did.

Based on this explorative investigation, I have the following two recommendations. Firstly, a phylogenetic tree should be produced for 16S rRNA gene amplicon datasets. It should be a standard procedure to visualise and inspect this phylogenetic tree to identify if erroneous sequences have been classified as bacteria. Inspecting the tree is a valuable practice even though phylogenetic based analysis is not performed to enhance the quality of the dataset. Additionally, it can be advantageous to include possible contaminant host- or ecosystem-related 16/18S rRNA

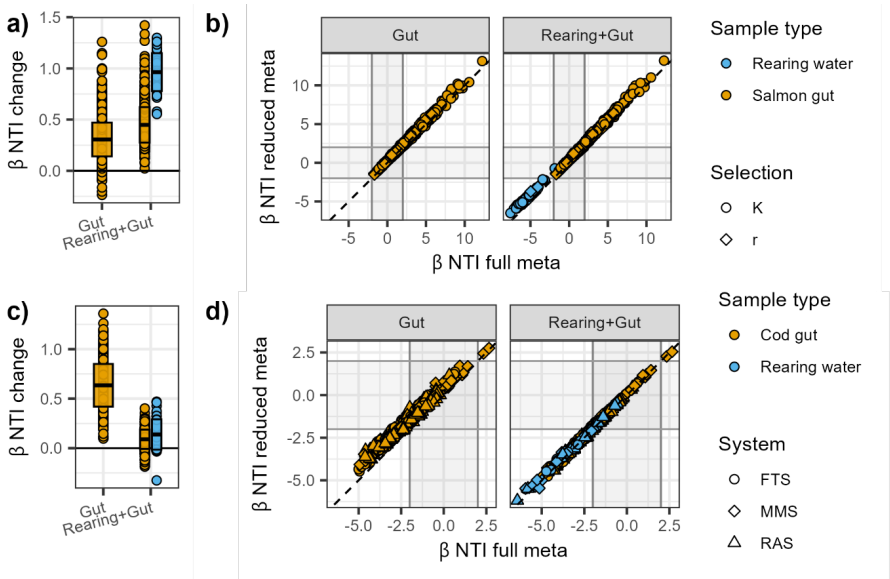


Figure 9: Estimated change in βNTI when the metacommunity was reduced. The change in βNTI for all comparisons of gut and rearing water samples was calculated as the difference in βNTI between the reduced and the original metacommunity in a) Dataset 1 and c) Dataset 2. The box plot indicates the mean change in βNTI ± standard deviation. b and d) Comparison of βNTI in the reduced and full metacommunity in b) Dataset 1 and d) Dataset 2. The dashed line indicates a 1:1 ratio. Values within the grey areas are within 2 standard deviations of the null model. Only sample comparisons within the same sample type, sampling day and sampling water selection regime are shown.

gene sequences in the reference dataset used in the bioinformatical pipelines to minimize the production of erroneous sequences. Secondly, when observed NTI or βNTI are close to the null-model estimates, the analysis should be done with various metacommunity sizes to evaluate the robustness of the community assembly conclusions. It should also be clearly stated what the metacommunity is based on when describing the analytical approach as there is a tendency that the more comparisons are deemed deterministic when the metacommunity is more diverse.

In conclusion, this exploratory data analysis highlights some limitations of the null models by emphasizing the impact of the phylogenetic tree on estimated NTI and βNTI, underscoring the importance of producing accurate trees and considering metacommunity size for robust community assembly conclusions.

References

Legendre, P., & Cáceres, M. De. (2013). Beta diversity as the variance of community data: dissimilarity coefficients and partitioning. *Ecology Letters*, *16*(8), 951–963. https://doi.org/10.1111/ELE.12141

Lozupone, C., & Knight, R. (2005). UniFrac: A new phylogenetic method for comparing microbial communities. *Applied and Environmental Microbiology*, *71*(12), 8228–8235. https://doi.org/10.1128/AEM.71.12.8228-8235.2005/ASSET/9AE40EDF-8FF9-4184-8DD7-367105439C23/ASSETS/GRAPHIC/ZAM0120562270003.JPEG

Mathisen, A. J. H. (2019). *Effekt av vertsgenetikk og mikrobiell vannkvalitet på kononiseringen av lakseyngel* [Master thesis, NTNU]. https://ntnuopen.ntnu.no/ntnu-xmlui/handle/11250/2630126

McMurdie, P. J., & Holmes, S. (2013). phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLOS ONE*, *8*(4), e61217. https://doi.org/10.1371/journal.pone.0061217

Nemergut, D. R., Schmidt, S. K., Fukami, T., O'Neill, S. P., Bilinski, T. M., Stanish, L. F., Knelman, J. E., Darcy, J. L., Lynch, R. C., Wickey, P., & Ferrenberg, S. (2013). Patterns and processes of microbial community assembly. *Microbiology and Molecular Biology Reviews : MMBR*, *77*(3), 342–356. https://doi.org/10.1128/MMBR.00051-12

Oksanen, J., Simpson, G., Blanchet, G., Kindt, R., Legendre, P., Minchin, P., O'Hara, R., Solymos, P., & et al. (2022). *vegan: Community Ecology Package* (R package version 2.6-4).

Stegen, J. C., Lin, X., Fredrickson, J. K., Chen, X., Kennedy, D. W., Murray, C. J., Rockhold, M. L., & Konopka, A. (2013). Quantifying community assembly processes and identifying features that impose them. *The ISME Journal*, *7*(11), 2069–2079. https://doi.org/10.1038/ismej.2013.93

Stegen, J. C., Lin, X., Konopka, A. E., & Fredrickson, J. K. (2012). Stochastic and deterministic assembly processes in subsurface microbial communities. *The ISME Journal*, *6*(9), 1653–1664. https://doi.org/10.1038/ismej.2012.22

Vellend, M. (2010). Conceptual synthesis in community ecology. *The Quarterly Review of Biology*, *85*(2), 183–206.

Vellend, M. (2016). *The Theory of Ecological Communities* (Vol. 57). Princeton University Press. https://press.princeton.edu/books/hardcover/9780691164847/the-theory-of-ecological-communities-mpb-57

Vestrum, R. I., Attramadal, K. J. K., Vadstein, O., Gundersen, M. S., & Bakke, I. (2020). Bacterial community assembly in Atlantic cod larvae (Gadus morhua): contributions of ecological processes and metacommunity structure. *FEMS Microbiology Ecology*, *96*(9), fiaa163. https://doi.org/https://doi.org/10.1093/femsec/fiaa163

Webb, C. O., Ackerly, D. D., McPeek, M. A., & Donoghue, M. J. (2002). Phylogenies and Community Ecology. *Annual Review of Ecology and Systematics*, *33*, 475–505. http://www.jstor.org/stable/3069271

Zhou, J., & Ning, D. (2017). Stochastic Community Assembly: Does It Matter in Microbial Ecology? *Microbiology and Molecular Biology Reviews*, *81*(4), e00002-17. https://doi.org/10.1128/MMBR.00002-17

Supplementary figures and tables:

Supplementary Table 1: The taxonomic identity of the top BLAST matches with highest similarity to the outlier sequences in Dataset 1. A maximum of five matches and only unique taxonomies are presented for each ASV.

ASV	Taxonomy
Zotu348	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Protacanthopterygii ; Salmoniformes; Salmonidae; Salmoninae; Salmo
Zotu348	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Protacanthopterygii ; Salmoniformes; Salmonidae; Coregoninae; Coregonus
Zotu530	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Protacanthopterygii ; Salmoniformes; Salmonidae; Salmoninae; Salmo
Zotu592	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Protacanthopterygii ; Salmoniformes; Salmonidae; Coregoninae; Coregonus
Zotu592	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Protacanthopterygii ; Salmoniformes; Salmonidae; Salmoninae; Salmo
Zotu614	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Protacanthopterygii ; Salmoniformes; Salmonidae; Salmoninae; Salmo
Zotu614	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Protacanthopterygii ; Salmoniformes; Salmonidae; Coregoninae; Coregonus
Zotu729	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Protacanthopterygii ; Salmoniformes; Salmonidae; Salmoninae; Salmo
Zotu729	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Neoteleostei; Acanthomorphata; Eupercaria; Perciformes; Notothenioidei; Channichthyidae; Pseudochaenichthys
Zotu729	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Neoteleostei; Acanthomorphata; Eupercaria; Perciformes; Notothenioidei; Bovichtidae; Cottoperca
Zotu729	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Protacanthopterygii ; Argentinidae; Argentina
Zotu910	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Protacanthopterygii ; Salmoniformes; Salmonidae; Salmoninae; Salmo
Zotu910	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Protacanthopterygii ; Salmoniformes; Salmonidae; Coregoninae; Coregonus
Zotu910	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Protacanthopterygii ; Salmoniformes; Salmonidae; Salmoninae; Salvelinus
Zotu969	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Protacanthopterygii ; Salmoniformes; Salmonidae; Salmoninae; Salmo
Zotu969	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Protacanthopterygii ; Salmoniformes; Salmonidae; Coregoninae; Coregonus
Zotu1008	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Protacanthopterygii ; Salmoniformes; Salmonidae; Salmoninae; Salmo
Zotu1008	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Protacanthopterygii ; Salmoniformes; Salmonidae; Coregoninae; Coregonus
Zotu1015	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Protacanthopterygii ; Salmoniformes; Salmonidae; Coregoninae; Coregonus
Zotu1015	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Protacanthopterygii ; Salmoniformes; Salmonidae; Salmoninae; Salmo
Zotu1020	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Protacanthopterygii ; Salmoniformes; Salmonidae; Salmoninae; Salmo
Zotu1096	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Protacanthopterygii ; Salmoniformes; Salmonidae; Salmoninae; Salmo
Zotu1177	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Protacanthopterygii ; Salmoniformes; Salmonidae; Salmoninae; Salmo
Zotu1177	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Protacanthopterygii ; Salmoniformes; Salmonidae; Coregoninae; Coregonus
Zotu1266	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Protacanthopterygii ; Salmoniformes; Salmonidae; Salmoninae; Salmo
Zotu1266	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Protacanthopterygii ; Salmoniformes; Salmonidae; Salmoninae; Oncorhynchus
Zotu1334	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Protacanthopterygii ; Salmoniformes; Salmonidae; Salmoninae; Salmo

Zotu1349	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Protacanthopterygii ; Salmoniformes; Salmonidae; Salmoninae; Salmo
Zotu1349	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Protacanthopterygii ; Salmoniformes; Salmonidae; Salmoninae; Oncorhynchus
Zotu1404	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Protacanthopterygii ; Salmoniformes; Salmonidae; Salmoninae; Salmo
Zotu1404	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Protacanthopterygii ; Salmoniformes; Salmonidae; Coregoninae; Coregonus
Zotu1404	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Protacanthopterygii ; Salmoniformes; Salmonidae; Salmoninae; Oncorhynchus
Zotu1462	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Protacanthopterygii ; Salmoniformes; Salmonidae; Salmoninae; Salmo
Zotu1654	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Protacanthopterygii ; Salmoniformes; Salmonidae; Salmoninae; Salmo
Zotu1654	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Protacanthopterygii ; Salmoniformes; Salmonidae; Salmoninae; Salvelinus
Zotu1742	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Protacanthopterygii ; Salmoniformes; Salmonidae; Salmoninae; Salmo
Zotu1742	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Protacanthopterygii ; Salmoniformes; Salmonidae; Coregoninae; Coregonus
Zotu1751	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Protacanthopterygii ; Salmoniformes; Salmonidae; Salmoninae; Salmo
Zotu1751	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Protacanthopterygii ; Salmoniformes; Salmonidae; Coregoninae; Coregonus
Zotu1752	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Protacanthopterygii ; Salmoniformes; Salmonidae; Salmoninae; Salmo
Zotu1752	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Protacanthopterygii ; Salmoniformes; Salmonidae; Coregoninae; Coregonus
Zotu1766	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Protacanthopterygii ; Salmoniformes; Salmonidae; Salmoninae; Salmo
Zotu1766	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Protacanthopterygii ; Salmoniformes; Salmonidae; Salmoninae; Salvelinus
Zotu1766	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Protacanthopterygii ; Salmoniformes; Salmonidae; Coregoninae; Coregonus
Zotu1801	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Protacanthopterygii ; Salmoniformes; Salmonidae; Salmoninae; Salmo
Zotu1801	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Protacanthopterygii ; Salmoniformes; Salmonidae; Salmoninae; Oncorhynchus
Zotu1824	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Protacanthopterygii ; Salmoniformes; Salmonidae; Salmoninae; Salmo
Zotu1824	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Protacanthopterygii ; Salmoniformes; Salmonidae; Salmoninae; Oncorhynchus
Zotu1863	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Protacanthopterygii ; Salmoniformes; Salmonidae; Salmoninae; Salmo
Zotu1863	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Neoteleostei; Acant homorphata; Eupercaria; Perciformes; Notothenioidei; Channichthyidae; Pseudochaenichthys
Zotu1863	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Neoteleostei; Acant homorphata; Carangaria; Pleuronectiformes; Pleuronectoidei; Pleuronectidae; Pleuronectes
Zotu1863	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Neoteleostei; Acant homorphata; Eupercaria; Perciformes; Notothenioidei; Bovichtidae; Cottoperca
Zotu1896	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Protacanthopterygii ; Salmoniformes; Salmonidae; Salmoninae; Salmo
Zotu1896	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Protacanthopterygii ; Salmoniformes; Salmonidae; Coregoninae; Coregonus
Zotu2052	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Protacanthopterygii ; Salmoniformes; Salmonidae; Salmoninae; Salmo
Zotu2052	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Protacanthopterygii ; Salmoniformes; Salmonidae; Coregoninae; Coregonus
Zotu2052	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cyprin iformes; Danionidae; Danioninae; Danio
Zotu2361	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Protacanthopterygii ; Salmoniformes; Salmonidae; Salmoninae; Salmo
Zotu2361	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Protacanthopterygii ; Salmoniformes; Salmonidae; Coregoninae; Coregonus

An exploration of the impact of phylogenetic tree structure on NTI and βNTI estimates of community assembly.

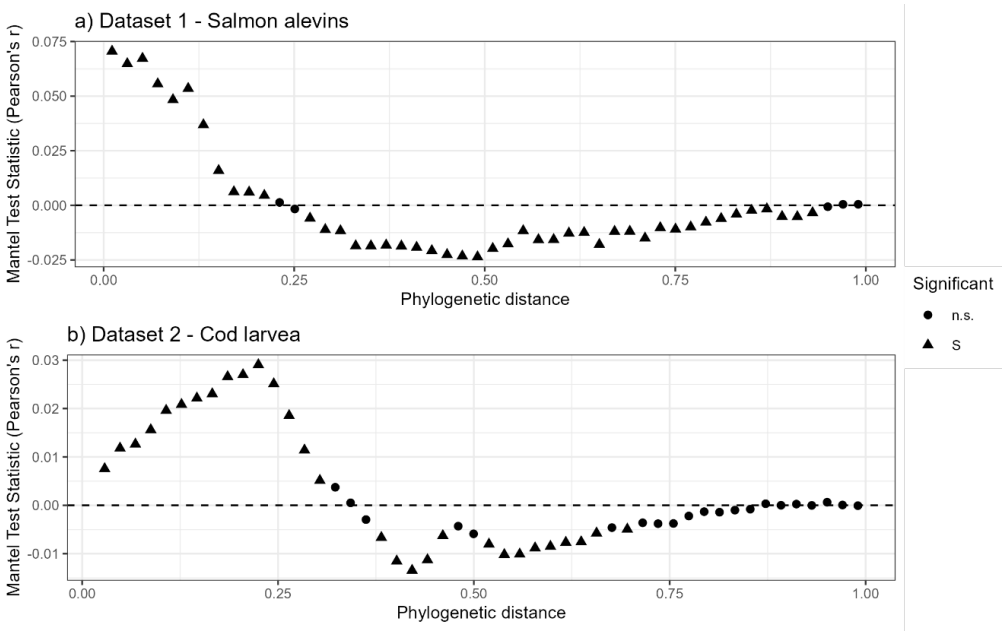
Zotu2446	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Protacanthopterygii ; Salmoniformes; Salmonidae; Salmoninae; Salmo
Zotu2448	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Protacanthopterygii ; Salmoniformes; Salmonidae; Salmoninae; Salmo
Zotu2448	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Protacanthopterygii ; Salmoniformes; Salmonidae; Salmoninae; Oncorhynchus
Zotu2765	Bacteria; Pseudomonadota; Gammaproteobacteria; Pseudomonadales; Pseudomonadaceae; Pseudomonas - Assumed to be wrongly classified due to highly diverging phylogenetic distance
Zotu2769	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Protacanthopterygii ; Salmoniformes; Salmonidae; Salmoninae; Salmo
Zotu2769	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Protacanthopterygii ; Salmoniformes; Salmonidae; Coregoninae; Coregonus
Zotu2821	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Protacanthopterygii ; Salmoniformes; Salmonidae; Salmoninae; Salmo
Zotu2821	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Protacanthopterygii ; Salmoniformes; Salmonidae; Coregoninae; Coregonus
Zotu2935	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Protacanthopterygii ; Salmoniformes; Salmonidae; Salmoninae; Salmo
Zotu3167	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Protacanthopterygii ; Salmoniformes; Salmonidae; Salmoninae; Salmo
Zotu3250	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Protacanthopterygii ; Salmoniformes; Salmonidae; Salmoninae; Salmo
Zotu3250	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Protacanthopterygii ; Salmoniformes; Salmonidae; Coregoninae; Coregonus
Zotu3285	Bacteria; Pseudomonadota; Betaproteobacteria; Burkholderiales; Oxalobacteraceae; Undibacterium - Assumed to be wrongly classified due to highly diverging phylogenetic distance
Zotu3285	Bacteria; environmental samples

Supplementary Table 2: The taxonomic identity of the top BLAST matches with highest similarity to the outlier sequences in Dataset 2. A maximum of five matches and only unique taxonomies are presented for each ASV.

OTU	Taxonomy
otu228	Eukaryota; Fungi; Dikarya; Ascomycota; Pezizomycotina; Dothideomycetes; Pleosporomycetidae; Pleosporales; Delitschiacea e; Delitschia
otu228	Eukaryota; Fungi; Dikarya; Ascomycota; Pezizomycotina; Dothideomycetes; Dothideomycetes incertae sedis; Zopfiaceae; Zop fia
otu228	Eukaryota; Fungi; Dikarya; Ascomycota; Pezizomycotina; Dothideomycetes; Pleosporomycetidae; Mytilinidiales; Mytilinidiace ae; Lophium
otu473	Eukaryota; Sar; Stramenopiles; Bigyra; Opalozoa; Bicosoecida; Caecitellus
otu1022	no BLAST matches
otu1245	no BLAST matches
otu1264	Eukaryota; Fungi; Dikarya; Ascomycota; Pezizomycotina; Sordariomycetes; Hypocreomycetidae; Hypocreales; Nectriaceae; Fu sarium; Fusarium oxysporum species complex
otu1264	Eukaryota; Fungi; Dikarya; Ascomycota; Pezizomycotina; Leotiomycetes; Helotiales; Dermateaceae; Phlyctema
otu1264	Eukaryota; Fungi; Dikarya; Ascomycota; Pezizomycotina; Leotiomycetes; Erysiphales; Erysiphaceae; Podosphaera
otu1264	Eukaryota; Fungi; Dikarya; Ascomycota; Pezizomycotina; Sordariomycetes; Hypocreomycetidae; Hypocreales; Sarocladiaceae; Sarocladium
otu1264	Eukaryota; Fungi; Dikarya; Ascomycota; Pezizomycotina; Sordariomycetes; Hypocreomycetidae; Hypocreales; Nectriaceae; Fu sarium
otu1421	Eukaryota; Sar; Stramenopiles; Bigyra; Opalozoa; Bicosoecida; Caecitellus
otu1639	Eukaryota; Sar; Stramenopiles; Bigyra; Opalozoa; Bicosoecida; Bicosoecida incertae sedis; Bilabrum
otu1639	Eukaryota; environmental samples
otu1837	Eukaryota; Sar; Stramenopiles; Bigyra; Opalozoa; Bicosoecida; Caecitellus
otu1837	Eukaryota; environmental samples
otu1887	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Neoteleostei; Acan thomorphata; Zeiogadaria; Gadariae; Gadiformes; Gadoidei; Gadidae; Gadus
otu1945	Eukaryota; environmental samples
otu1945	Eukaryota; Sar; Stramenopiles; Bigyra; Opalozoa; Nanomonadea; Incisomonas
otu2101	Eukaryota; Fungi; Dikarya; Ascomycota; Pezizomycotina; Dothideomycetes; Pleosporomycetidae; Pleosporales; Pleosporinea e; Pleosporaceae; Alternaria; Alternaria sect. Infectoriae
otu2101	Eukaryota; Fungi; Dikarya; Ascomycota; Pezizomycotina; Dothideomycetes; Pleosporomycetidae; Pleosporales; Pleosporinea e; Pleosporaceae; Alternaria; Alternaria sect. Porri
otu2285	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Neoteleostei; Acan thomorphata; Zeiogadaria; Gadariae; Gadiformes; Gadoidei; Gadidae; Gadus
otu2312	Eukaryota; Fungi
otu2312	Eukaryota; Fungi; Dikarya; Ascomycota; Pezizomycotina; Dothideomycetes; Pleosporomycetidae; Pleosporales; Pleosporinea e; Didymellaceae; Phoma
otu2312	Eukaryota; Fungi; Dikarya; Ascomycota; Pezizomycotina; Dothideomycetes; Pleosporomycetidae; Pleosporales; Pleosporinea e; Didymellaceae; Ascochyta
otu2365	Eukaryota; Sar; Stramenopiles; environmental samples
otu2365	Eukaryota
otu2365	Eukaryota; Sar; Stramenopiles; Bigyra; Labyrinthulomycetes; Thraustochytrida; Thraustochytriaceae
otu2365	Eukaryota; environmental samples
otu2487	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Neoteleostei; Acan thomorphata; Zeiogadaria; Gadariae; Gadiformes; Gadoidei; Gadidae; Gadus
otu2487	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Chondrostei; Acipenseriformes; Acipenser idae; Acipenser

otu2541	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Neoteleostei; Acanthomorphata; Zeiogadaria; Gadariae; Gadiformes; Gadoidei; Gadidae; Gadus
otu2704	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Neoteleostei; Acanthomorphata; Zeiogadaria; Gadariae; Gadiformes; Gadoidei; Gadidae; Gadus
otu3087	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Neoteleostei; Acanthomorphata; Zeiogadaria; Gadariae; Gadiformes; Gadoidei; Gadidae; Gadus
otu3120	Bacteria; environmental samples
otu3120	unclassified sequences; environmental samples
otu3120	Archaea; environmental samples
otu3159	Eukaryota; Fungi; Dikarya; Ascomycota; Pezizomycotina; Sordariomycetes; Hypocreomycetidae; Hypocreales; Nectriaceae; Fusarium; Fusarium solani species complex
otu3159	Eukaryota; Fungi; Dikarya; Ascomycota; Pezizomycotina; Dothideomycetes; Pleosporomycetidae; Pleosporales; Pleosporinea e; Didymellaceae; Phoma
otu3159	Eukaryota; Fungi; Dikarya; Ascomycota; Pezizomycotina; Dothideomycetes; Pleosporomycetidae; Pleosporales; Pleosporinea e; Didymellaceae; Didymella
otu3159	Eukaryota; Fungi
otu3227	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Neoteleostei; Acanthomorphata; Holocentriformes; Holocentridae; Myripristis
otu3412	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Neoteleostei; Acanthomorphata; Zeiogadaria; Gadariae; Gadiformes; Gadoidei; Gadidae; Gadus
otu3803	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Neoteleostei; Acanthomorphata; Zeiogadaria; Gadariae; Gadiformes; Gadoidei; Gadidae; Gadus
otu3806	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Neoteleostei; Acanthomorphata; Zeiogadaria; Gadariae; Gadiformes; Gadoidei; Gadidae; Gadus
otu3806	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Neoteleostei; Acanthomorphata; Ovalentaria; Pomacentridae; Acanthochromis
otu3806	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Neoteleostei; Acanthomorphata; Ovalentaria; Blenniimorphae; Blenniiformes; Blennioidei; Blenniidae; Salariinae; Salarias
otu4365	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Neoteleostei; Acanthomorphata; Zeiogadaria; Gadariae; Gadiformes; Gadoidei; Gadidae; Gadus
otu4369	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Neoteleostei; Acanthomorphata; Zeiogadaria; Gadariae; Gadiformes; Gadoidei; Gadidae; Gadus
otu5643	no BLAST matches
otu7224	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Neoteleostei; Acanthomorphata; Zeiogadaria; Gadariae; Gadiformes; Gadoidei; Gadidae; Gadus
otu567	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Neoteleostei; Acanthomorphata; Zeiogadaria; Gadariae; Gadiformes; Gadoidei; Gadidae; Gadus
otu1841	unclassified sequences; environmental samples
otu1841	Archaea; Euryarchaeota; environmental samples
otu1841	Archaea; environmental samples
otu2033	Eukaryota; Viridiplantae; Chlorophyta; Mamiellophyceae; Mamiellales; Bathycoccaceae; Bathycoccus
otu2033	Bacteria; environmental samples
otu2033	Bacteria; Pseudomonadota; Gammaproteobacteria; Pseudomonadales; Pseudomonadaceae; Pseudomonas; environmental samples
otu2033	Bacteria; Pseudomonadota; Alphaproteobacteria; Rickettsiales; Anaplasmataceae; Anaplasma
otu2568	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Neoteleostei; Acanthomorphata; Zeiogadaria; Gadariae; Gadiformes; Gadoidei; Gadidae; Gadus
otu3349	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Neoteleostei; Acanthomorphata; Zeiogadaria; Gadariae; Gadiformes; Gadoidei; Gadidae; Gadus
otu4489	no BLAST matches_check up
otu4831	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Neoteleostei; Acanthomorphata; Zeiogadaria; Gadariae; Gadiformes; Gadoidei; Gadidae; Gadus
otu4831	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Neoteleostei; Acanthomorphata; Eupercaria; Spariformes; Sparidae; Acanthopagrus

otu1295	no BLAST matches
otu3413	no BLAST matches
otu4047	no BLAST matches
otu365	Eukaryota; Discoba; Jakobida; Histionina; Jakobidae; Jakoba
otu365	Bacteria; environmental samples
otu2254	Eukaryota; Discoba; Jakobida; Histionina; Seculamonas
otu3237	Eukaryota; Amoebozoa; Discosea; Longamoebia; Centramoebida; Acanthamoebidae; Acanthamoeba
otu6048	Bacteria; environmental samples
otu6048	Eukaryota; Rhodophyta; Florideophyceae; Nemaliophycidae; Batrachospermales; Lemaneaceae; Paralemanea
otu6048	unclassified sequences; environmental samples



Supplementary Figure 1: Phylogenetic Mantel correlogram indicating significant phylogenetic signals across short phylogenetic distances in a) Dataset 1 and b) Dataset 2. Significance indicates whether between ASV/OTU niche is correlated to differences between ASV/OTU phylogenetic differences across a phylogenetic distance (s=significant, $p<0.05$, n.s. = not significant). The Mantel correlogram indicates that ASVs/OTUs inhabiting the same niches have close phylogenetic distances, and thus there is a phylogenetic signal in the dataset.

quality_control_sequences.R

by Madeleine Gundersen, 2023-06-08

```
#the function takes in a data frame containing all taxa outliers.
#For each taxa (Query) it checks if its a bacteria
# df is the dataframe with the outlier taxa and taxonomic info
quality_control <- function(df) {
  Query = unique(df$Query)
  # Initiate a result column
  quality_control_df = data.frame(Query = Query, Decition = 0)

  cat("\n")
  cat("\033[31mPress ESC to cancel\033[0m\n")
  Sys.sleep(1)
  cat("\033[31mFor each taxa decide if its an outlier\033[0m\n")
  Sys.sleep(1)
  cat("\033[32mGood luck! ~ Madeleine :-) \033[0m\n")
  cat("\n")
  Sys.sleep(1)

  # Iterate for all Queries
  for (i in seq_along(Query)) {
  Query_sub = Query[i]
  # filter for rows with the given query
  query_df <- subset(df, Query == Query_sub)
  query_df = query_df %>% dplyr::mutate(taxonomy = paste0(substr(taxonomy, start = 1, stop = 5
0), "..."))
  # subset dataframe
  unique_query_df = query_df[, c("Query", "Organism", "taxonomy")]

  # check if all unique rows have "Bacteria" as the first 8 characters of the taxonomy
  tax_unique <- unique(substr(unique_query_df$taxonomy, start = 1, stop = 8))
  if (length(tax_unique) == 1 && grepl("Bacteria", tax_unique)) {
    cat(sprintf("Query: %s is a bacteria. Keeping it.\n", Query_sub))
    cat("\n")
    quality_control_df[i,2] = "KEEP"

  } else {
    # display the unique rows to the user
    cat(sprintf("Query: %s\n", Query_sub))
    print(unique_query_df)

    # ask the user if they want to keep the query
    keep_query <- readline(sprintf("Do you want to keep query %s? (yes or no) ", Query_sub))

    # return the query and whether or not to keep it
    if (tolower(keep_query) == "yes" || tolower(keep_query) == "y") {
      cat(sprintf("\033[32mOk - then it's decided %s is a bacteria.\n\033[0m", Query_sub))
      Sys.sleep(1)
```

```
      cat("\n")
      cat("\n")
      quality_control_df[i,2] = "KEEP"

      } else {
      cat(sprintf("\033[31mOk - let's get rid of this intruder ;) %s has been eliminated!\033[
0m\n", Query_sub))
      Sys.sleep(1)
      cat("\n")
      cat("\n")
      quality_control_df[i,2] = "OUTLIER"
      } }
  Sys.sleep(0.5)
  }
  return(quality_control_df) }
```

Plot phylogenetic tree and identify outliers

Madeleine J.S. Gundersen

2023-05-09

This R-notebook document shows how a user can investigate their phylogenetic tree and identify outliers. The outline is as follows

1. Input your phyloseq object and plot is with cirlular layout using ggtree package.
2. If outliers are present identify them
3. Give the script your fasta file and it will make a new fasta file containing only the “suspicious” sequences
4. Input this new fasta file to BLAST and download the XML report for all samples.
5. Input the XML report back into R
6. In this script the top 5 BLAST matches per OTU/ASV is kept. But you can modify this to your preference.
7. Match the accession number for each match to taxonomic information from NCBI. If all 5 matches are “Bacteria” we keep the ASV. If they are not bacteria, the program will display the BLAST information to you and you have to decide if the OTU/ASV should be removed.
8. You are given a vector containing the biased sequences and can make a new phyloseq object without these biased ASVs/OTUs.
9. Plot the tree again and see if you need to repeat.

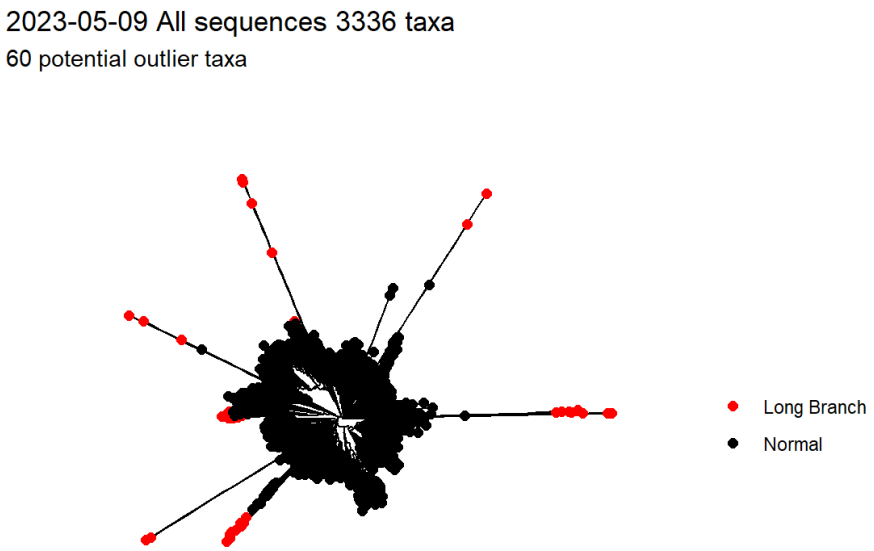
```
# load packages
library(phyloseq)
library(ggtree)
library(ggplot2)
library(tibble)
library(dplyr)
library(xml2)
library(rentrez)
library(kableExtra)
```

```
# Set filepaths to your data and for saving figures and resuls
filepath= "the_path_to_your_folder"
# e.g. "C:/Users/madel/Project_Cod/R_analysis
filepath_results = paste0(filepath,"/figures/", Sys.Date(), "_")

#load the phyloseq object
phyloseq_all_sequences = readRDS("your_phyloseq_object.rds")
fasta_file_experiment = "your_project_fasta_file.fasta"
```

In this document an experiment containing outlier sequences was used to exemplify the code and method.

Inspect the phylogenetic tree and identify outliers

```
# Name title to be displayed in plot
dataset_title = "All sequences"

# Display the tree in a circular layout using the ggtree package
ggtree_obj = ggtree::ggtree(phyloseq_all_sequences) +
  ggtree::layout_circular()

# Identify taxa with long branches
branch_cutoff_outlier = 0.32#adjust if needed
long_branch_taxa = ggtree_obj$data$label[ggtree_obj$data$branch > branch_cutoff_outlier]

# Find number and names of outlier taxa
long_branch_taxa_df = data.frame(outliers = long_branch_taxa) %>% unique() %>% na.omit()

# Highlight taxa with long branch lengths in red
# Display the updated tree
ggtree_obj +
  ggtree::geom_tippoint(aes(color = ifelse(label %in% long_branch_taxa, "Long Branch", "Norma
l")),
                        size = 2) +
  ggplot2::scale_color_manual(values = c("Normal" = "black", "Long Branch" = "red")) +
  ggplot2::theme(legend.position = "right",
                 legend.title = element_blank()) +
  ggplot2::ggtitle(label = paste0(Sys.Date(), " ", dataset_title, " ", ntaxa(phyloseq_all_sequen
ces), " taxa"),
                   subtitle = paste0(nrow(long_branch_taxa_df), " potential outlier taxa"))
```




```
# Save the plot if desired
plotname = "Phylogenetic_tree_all_sequences_with_outliers"
#ggplot2::ggsave(filename = paste0(filepath_results, plotname, ".png"), width = 10, height = 10)
```

```
# You can also filter out outliers using other indicators

## Based on taxonomy

#unclassified_bacteria = ggtree_obj$data$label[ggtree_obj$data$Phylum == "uncl_d_Bacteria"]
```

Export outlier sequences in fasta format

```
seq_data <- Biostrings::readDNAStringSet(fasta_file_experiment)
```

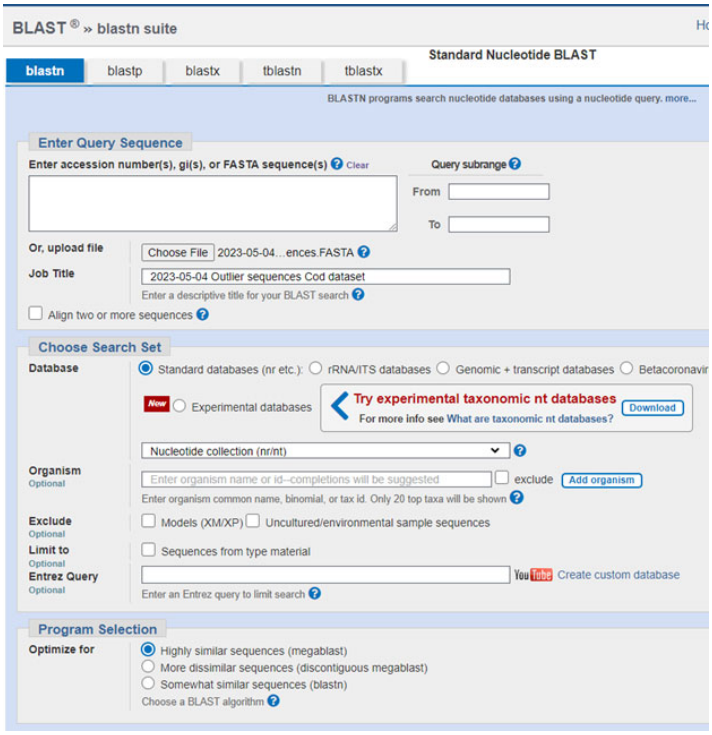
```
# Filter sequencing data to only contain outliers
outlier_seq_data <- seq_data[names(seq_data) %in% long_branch_taxa_df$outliers, ]

#export outlier sequences in fasta format for input to BLAST
fasta_file_name = "Outlier_sequences"
Biostrings::writeXStringSet(x = outlier_seq_data,
                            filepath = paste0(filepath_results,
                                              fasta_file_name, "_",
                                              length(outlier_seq_data), "sequences.fasta"))
```

```
# Create data frame with information for outlier taxa. Can come if handy if you want to check up
specific taxa
taxa_df <- data.frame(taxaID = rownames(tax_table(phyloseq_all_sequences)),
                      tax_table(phyloseq_all_sequences)) %>%
  dplyr::filter(taxaID %in% long_branch_taxa_df$outliers)

sequence_df =
data.frame(taxaID = names(outlier_seq_data),
  Sequence = as.character(outlier_seq_data))

outlier_info_df = left_join(taxa_df, sequence_df, by = "taxaID")

#save as a csv
write.csv(x = outlier_info_df, file = paste0(filepath_results, "outlier_phylogenetic_tree_ID_seq
uence.csv"))
```

Nucleotide BLAST the outlier sequences

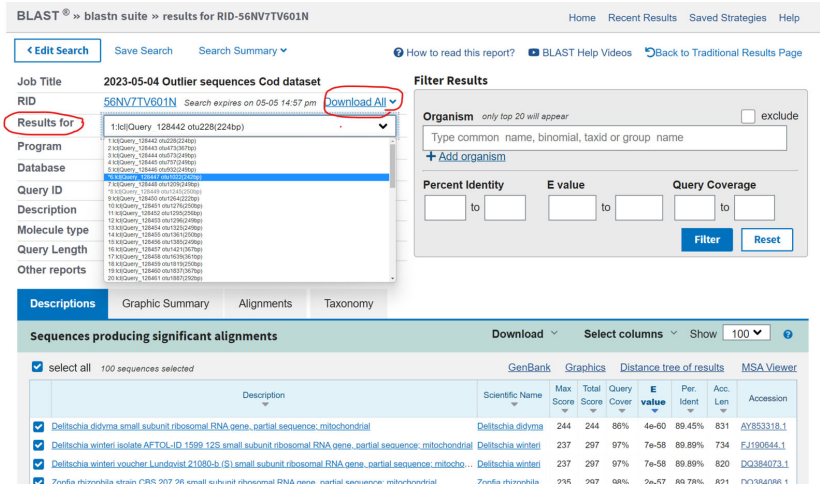
Go to https://blast.ncbi.nlm.nih.gov/Blast.cgi (https://blast.ncbi.nlm.nih.gov/Blast.cgi) and choose “Nucleotide BLAST”.

Upload the fasta file continaing the potential outlier sequences and BLAST against Nucleotide collection (nr/nt). Try first using standard algorithm parameters.



BLAST search window

When the search is finished you will get up a result window. In this case I uploaded 60 sequences, and should get 60 Query summaries back. Check this by clicking “Results for” button. Then click “Download All” and select “XML”. The XML file should now be in your download folder. Transfer this file to your project folder and rename if wanted.



Analyse BLAST results

Load XML files and format information to data frame structure

```
# Load the XML file (one XML file per BLAST)
# set the directory where the XML files are located
xml_dir <- paste0(filepath)
# get the filenames of the XML files
xml_files <- list.files(xml_dir, pattern = "^.*\\.xml$")
# read the XML files into a list
xml_files = paste0(filepath, xml_files)
xml_list <- lapply(xml_files, xml2::read_xml)
```

You can inspect the XML file by clinking the file in the “Files” window to get a grasp of the structure. Here we are saving the description, accession number , and E-value

```
# Initialize empty data frame to store results.
results_df = data.frame(matrix(ncol = 4, nrow = 0))
colnames(results_df) = c("Query", "Accession", "Match", "E_Value")

for (list_position in 1:length(xml_list)) {
  # if you have more XML files this script will loop over them
  xmlfile = xml_list[[list_position]]

  # Get all query nodes (one node per taxa)
  query_nodes <-
    xml2::xml_find_all(xmlfile, "//BlastOutput_iterations/Iteration")

  # Loop through query nodes and extract information
  for (i in 1:length(query_nodes)) {
    # Get query name
    query_name = xml2::xml_find_first(query_nodes[i],
                    ".//Iteration_query-def") %>%
      xml2::xml_text()

    # Get top 5 matches
    hit_nodes <- xml_find_all(query_nodes[i], ".//Hit") %>%
      head(5) #choose desired matches to keep

    # Extract match information
    for (j in 1:length(hit_nodes)) {
      match_name = xml2::xml_find_first(hit_nodes[j], ".//Hit_def") %>% xml2::xml_text()
      accession_id = xml2::xml_find_first(hit_nodes[j], ".//Hit_accession") %>% xml2::xml_text()
      evalue = xml2::xml_find_first(hit_nodes[j], ".//Hsp_evalue") %>% xml2::xml_text()
      if (length(accession_id) == 0) {
        accession_id = "no BLAST matches_check up"
        match_name = "no BLAST matches_check up"
        evalue = "no BLAST matches_check up"}
    #Save information in result_df
      results_df = base::rbind(results_df,
                               data.frame(Query = query_name,
                                          Accession = accession_id,
                                          Match = match_name,
                                          E_Value = evalue  ))}}}
```

Query	Accession	Match	E_Value
otu228	AY853318	Delitschia didyma small subunit ribosomal RNA gene, partial sequence; mitochondrial	4.06238e-60
otu228	FJ190644	Delitschia winteri isolate AFTOL-ID 1599 12S small subunit ribosomal RNA gene, partial sequence; mitochondrial	6.79779e-58
otu228	DQ384073	Delitschia winteri voucher Lundqvist 21080-b (S) small subunit ribosomal RNA gene, partial sequence; mitochondrial	6.79779e-58
otu228	DQ384086	Zopfia rhizophila strain CBS 207.26 small subunit ribosomal RNA gene, partial sequence; mitochondrial	2.44492e-57

Query	Accession	Match	E_Value
otu228	KT225540	Lophium mytilinum isolate AFTOL-ID 1609 12S ribosomal RNA gene, partial sequence; mitochondrial	1.47146e-54
otu473	KY980331	Caecitellus paraparvulus isolate BH56_230 small subunit ribosomal RNA gene, partial sequence	0
otu473	KY980321	Caecitellus paraparvulus isolate BH56_212 small subunit ribosomal RNA gene, partial sequence	0
otu473	KY980317	Caecitellus paraparvulus isolate BH56_205 small subunit ribosomal RNA gene, partial sequence	0
otu473	KY980271	Caecitellus paraparvulus isolate BH56_108 small subunit ribosomal RNA gene, partial sequence	0
otu473	KY980264	Caecitellus paraparvulus isolate BH56_74 small subunit ribosomal RNA gene, partial sequence	0
otu573	KY124855	Uncultured bacterium clone A494_Bac213 16S ribosomal RNA gene, partial sequence	1.95201e-123
otu573	KP937587	Uncultured bacterium clone OTU95932_AL220_211882 16S ribosomal RNA gene, partial sequence	4.25535e-115
otu573	KP909788	Uncultured bacterium clone 4368878_AL223_1289544 16S ribosomal RNA gene, partial sequence	4.25535e-115
otu573	KR846765	Uncultured bacterium clone OTU_15481 16S ribosomal RNA gene, partial sequence	9.21121e-112
otu573	KU115717	Uncultured bacterium clone 19455 16S ribosomal RNA gene, partial sequence	9.21121e-112

Collect taxonomic information from NCBI

```
# create an empty dataframe to store taxonomy information
taxonomy_df <- data.frame(Organism = character(),
                          taxonomy = character(),
                          Accession = character(),
                          stringsAsFactors = FALSE)

# save the accession numbers in a dataframe and remove duplicate numbers
Accession_IDs = results_df %>% select(Accession) %>% unique()

# match the accession numbers with taxonomic information
for (i in 1:nrow(Accession_IDs)) {
  # skip if Accession is missing
  if (Accession_IDs$Accession[i] == "no BLAST matches_check up") {
    taxonomy_df <- rbind(taxonomy_df, data.frame(Organism = "no BLAST matches_check up",
                                                 taxonomy = "no BLAST matches_check up",
                                                 Accession = "no BLAST matches_check up",
                                                 stringsAsFactors = FALSE))
    next
  }
  # use efetch to retrieve the record from GenBank
  gb = rentrez::entrez_fetch(db = "nuccore", id = Accession_IDs$Accession[i], rettype = "XML")
  # convert the XML object to a list
  xml_list = XML::xmlToList(gb)
  # extract the required information
  organism = xml_list$GBSeq$GBSeq_organism
  taxonomy = xml_list$GBSeq$GBSeq_taxonomy
  accession = xml_list$GBSeq$`GBSeq_primary-accession`
  # create a data frame
  taxonomy_df = rbind(taxonomy_df, data.frame(Organism = organism,
                                              taxonomy = taxonomy,
                                              Accession = accession,
                                              stringsAsFactors = FALSE))}

# merge taxonomy information with results dataframe
outliers_with_accession = left_join(results_df, taxonomy_df, by = "Accession")
saveRDS(object = outliers_with_accession, file = paste0(filepath_results, "outliers_BLASTed_with
accession.RDS"))
```

```
outliers_with_accession = readRDS(paste0(filepath,"2023-05-04_outliers_BLASTed_withaccession.RD
S"))
# Check that all outliers have been analyed
# The numer of queries should be the same as possible outliers
outliers_with_accession %>% select(Query) %>% unique() %>% nrow()
```

```
## [1] 60
```

Quality control of outlier taxa

```
# Load quality_control script
source(file = paste0(filepath, "quality_control_sequences.R"))

# COPY the command below into the Console window and make a decition for each taxa
quality_control_decition = quality_control(df = outliers_with_accession)
outliers_with_accession_decition = dplyr::left_join(outliers_with_accession, quality_control_dec
ition, by = "Query")
saveRDS(outliers_with_accession_decition, file = paste0(filepath_results, "Quality_control_decit
ion.RDS"))
```

After running the script we now have a decition for each outlier taxa

```
outliers_with_accession_decition[1:15,] %>%
  kbl() %>%
  kable_paper("hover", full_width = F)
```

Query	Accession	Match	E_Value	Organism	taxonomy	Decition
otu228	AY853318	Delitschia didyma small subunit ribosomal RNA gene, partial sequence; mitochondrial	4.06238e-60	Delitschia didyma	Eukaryota; Fungi; Dikarya; Ascomycota; Pezizomycotina; Dothideomycetes; Pleosporomycetidae; Pleosporales; Delitschiaceae; Delitschia	OUTLIER
otu228	FJ190644	Delitschia winteri isolate AFTOL-ID 1599 12S small subunit ribosomal RNA gene, partial sequence; mitochondrial	6.79779e-58	Delitschia winteri	Eukaryota; Fungi; Dikarya; Ascomycota; Pezizomycotina; Dothideomycetes; Pleosporomycetidae; Pleosporales; Delitschiaceae; Delitschia	OUTLIER
otu228	DQ384073	Delitschia winteri voucher Lundqvist 21080-b (S) small subunit ribosomal RNA gene, partial sequence; mitochondrial	6.79779e-58	Delitschia winteri	Eukaryota; Fungi; Dikarya; Ascomycota; Pezizomycotina; Dothideomycetes; Pleosporomycetidae; Pleosporales; Delitschiaceae; Delitschia	OUTLIER
otu228	DQ384086	Zopfia rhizophila strain CBS 207.26 small subunit ribosomal RNA gene, partial sequence; mitochondrial	2.44492e-57	Zopfia rhizophila	Eukaryota; Fungi; Dikarya; Ascomycota; Pezizomycotina; Dothideomycetes; Dothideomycetes incertae sedis; Zopfiaceae; Zopfia	OUTLIER
otu228	KT225540	Lophium mytilinum isolate AFTOL-ID 1609 12S ribosomal RNA gene, partial sequence; mitochondrial	1.47146e-54	Lophium mytilinum	Eukaryota; Fungi; Dikarya; Ascomycota; Pezizomycotina; Dothideomycetes; Pleosporomycetidae; Mytilinidiales; Mytilinidiaceae; Lophium	OUTLIER
otu473	KY980331	Caecitellus paraparvulus isolate BH56_230 small subunit ribosomal RNA gene, partial sequence	0	Caecitellus paraparvulus	Eukaryota; Sar; Stramenopiles; Bigyra; Opalozoa; Bicosoecida; Caecitellus	OUTLIER

Query	Accession	Match	E_Value	Organism	taxonomy	Decition
otu473	KY980321	Caecitellus paraparvulus isolate BH56_212 small subunit ribosomal RNA gene, partial sequence	0	Caecitellus paraparvulus	Eukaryota; Sar; Stramenopiles; Bigyra; Opalozoa; Bicosoecida; Caecitellus	OUTLIER
otu473	KY980317	Caecitellus paraparvulus isolate BH56_205 small subunit ribosomal RNA gene, partial sequence	0	Caecitellus paraparvulus	Eukaryota; Sar; Stramenopiles; Bigyra; Opalozoa; Bicosoecida; Caecitellus	OUTLIER
otu473	KY980271	Caecitellus paraparvulus isolate BH56_108 small subunit ribosomal RNA gene, partial sequence	0	Caecitellus paraparvulus	Eukaryota; Sar; Stramenopiles; Bigyra; Opalozoa; Bicosoecida; Caecitellus	OUTLIER
otu473	KY980264	Caecitellus paraparvulus isolate BH56_74 small subunit ribosomal RNA gene, partial sequence	0	Caecitellus paraparvulus	Eukaryota; Sar; Stramenopiles; Bigyra; Opalozoa; Bicosoecida; Caecitellus	OUTLIER
otu573	KY124855	Uncultured bacterium clone A494_Bac213 16S ribosomal RNA gene, partial sequence	1.95201e-123	uncultured bacterium	Bacteria; environmental samples	KEEP
otu573	KP937587	Uncultured bacterium clone OTU95932_AL220_211882 16S ribosomal RNA gene, partial sequence	4.25535e-115	uncultured bacterium	Bacteria; environmental samples	KEEP
otu573	KP909788	Uncultured bacterium clone 4368878_AL223_1289544 16S ribosomal RNA gene, partial sequence	4.25535e-115	uncultured bacterium	Bacteria; environmental samples	KEEP
otu573	KR846765	Uncultured bacterium clone OTU_15481 16S ribosomal RNA gene, partial sequence	9.21121e-112	uncultured bacterium	Bacteria; environmental samples	KEEP
otu573	KU115717	Uncultured bacterium clone 19455 16S ribosomal RNA gene, partial sequence	9.21121e-112	uncultured bacterium	Bacteria; environmental samples	KEEP

```
outliers_remove = outliers_with_accession_decition %>% dplyr::filter(Decition == "OUTLIER") %>%
select(Query) %>% unique()
# Convert taxonomy table to data frame and add taxa id as a column
taxa_df <- data.frame(taxaID = rownames(tax_table(phyloseq_all_sequences)), tax_table(phyloseq_a
ll_sequences))
tax_table(phyloseq_all_sequences) = tax_table(as.matrix(taxa_df))

# Remove outliers from phyloseq object
quality_ps = subset_taxa(physeq = phyloseq_all_sequences, !taxaID %in% outliers_remove$Query)
quality_ps = filter_taxa(quality_ps, function(x) sum(x) > 0, prune = TRUE)
#saveRDS(quality_ps, file = paste0(filepath_results, "phyloseq_quality1_3308taxa.RDS") )
```

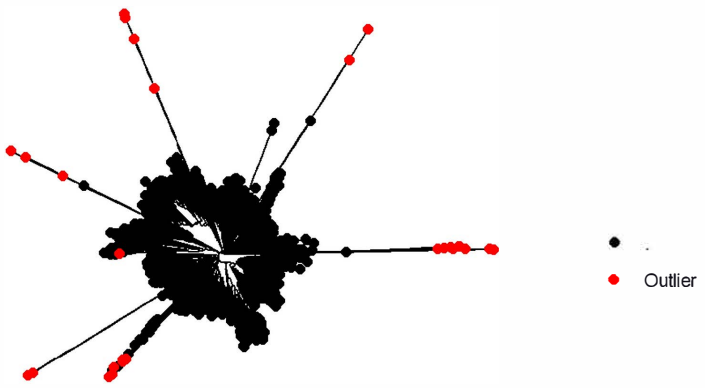
Plot new phylogenetic tree and filter out outliers

Finally, we remove the outliers from the phyloseq object and plot the new phylogenetic tree to evaluate if we need to repeat the process.

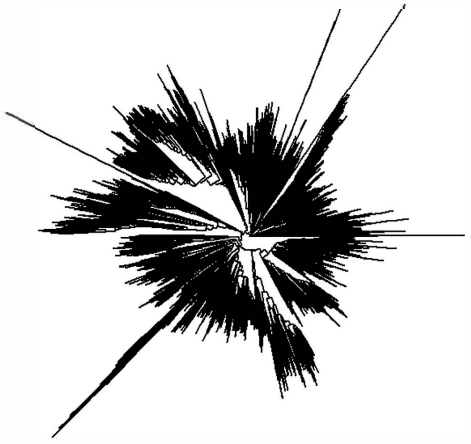
```
removed_tree=
ggtree(phyloseq_all_sequences) + layout_circular() +
  ggtree:: geom_tippoint(aes(color = ifelse(label %in% outliers_remove$Query, "Outlier", ".")),
size = 2) +
  ggplot2::scale_color_manual(values = c("." = "black", "Outlier" = "red")) +
  ggplot2::theme(legend.position = "right",
                 legend.title = element_blank()) +
  ggplot2::ggtitle(label = paste0(Sys.Date(), " Full dataset ", ntaxa(phyloseq_all_sequences), "
taxa"),
                   subtitle = paste0(nrow(outliers_remove), " outlier taxa"))
quality_tree =
ggtree(quality_ps) + layout_circular() +
  ggplot2::theme(legend.position = "right",
                 legend.title = element_blank()) +
  ggplot2::ggtitle(label = paste0(Sys.Date(), " Without outliers ", ntaxa(quality_ps), " taxa"))

ggpubr::ggarrange(removed_tree,quality_tree,nrow = 2)
```

2023-05-09 Full dataset 3336 taxa

28 outlier taxa



2023-05-09 Without outliers 3308 taxa



```
#ggsave(filename = paste0(filepath_results, #"removed_outliers_and_new_tree.png"), width = 6, he
ight = 10)
```

In this example it can be advisable to repeat the process as some taxa still are deviating from the phylogenetic tree.

ISBN 978-82-326-7504-3 (printed ver.)
ISBN 978-82-326-7503-6 (electronic ver.)
ISSN 1503-8181 (printed ver.)
ISSN 2703-8084 (online ver.)

NTNU

Norwegian University of
Science and Technology