Assembly and functioning of microbial communities along terrestrial resource gradients in boreal lake sediments

Chloé Shoshana Jessica Orland

Clare Hall

University of Cambridge

September 2018

This dissertation is submitted for the degree of Doctor of Philosophy.

Assembly and functioning of microbial communities along terrestrial resource gradients in boreal lake sediments

Chloé Shoshana Jessica Orland

Summary

Terrestrial inputs of organic matter contribute greatly to the functioning of aquatic ecosystems, subsidizing between 30-70% of secondary production. This contribution of terrestrial resources is especially important in boreal lakes that are largely nutrient-poor and thus more responsive to these additions. Yet the mechanisms underlying initial processing of terrestrial resources by microbial communities at the base of lake food webs remain poorly understood. With this in mind, this thesis aims to advance our understanding of lake sediment microbial community assembly and functioning along abiotic gradients, primarily reflecting variation in terrestrial organic matter inputs that are predicted to increase with future environmental change.

Chapter 1 reviews current knowledge on the terrestrial support of lake food webs and highlights gaps in understanding the factors influencing the microbial processing of terrestrial resources. It also provides an overview of metagenomics methods for microbial community analysis and their development over the course of the thesis. **Chapter 2** tests how much of ecosystem functioning is explained by microbial community structure relative to other ecosystem properties such as the present-day and past environment. Theory predicts that ecosystem functioning, here measured as CO_2 production, should increase with diversity, but the individual and interactive effects of other ecosystem properties on ecosystem functioning remain unresolved. **Chapter 3** further questions the importance of microbial diversity for ecosystem functioning by asking whether more diverse microbial communities stabilize ubiquitous functions like

CO_2 production and microbial abundances through time. It also aims to identify the biotic and abiotic mechanisms underlying positive diversity-stability relationships. **Chapter 4** then explores how microbial communities assemble and colonize sediments with varying types and amounts of terrestrial organic matter in three different lakes over a two-month period. Understanding how microbial communities change in relation to sediment and lake conditions can help predict downstream ecosystem functions. Finally, **Chapter 5** discusses the main findings of the thesis and ends with proposed avenues for future research.

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the Preface and specified in the text.

It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. I further state that no substantial part of my dissertation has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text.

It does not exceed the prescribed word limit of 60,000 for the Biology Degree Committee.

À mes grands-mères,

Mamie Genia et Mamie Marie

"A lake is the landscape's most beautiful and expressive feature. It is Earth's eye; looking into which the beholder measures the depth of his own nature."

~ Henry David Thoreau

"Perhaps the truth depends on a walk around the lake."

~ Wallace Stevens

Acknowledgments

I would first like to thank my supervisor Dr Andrew Tanentzap for giving me the opportunity to pursue my research interests and for being so generous with his time. Andrew, you somehow knew that I would complete this thesis even in my greatest moments of doubt and, importantly, you always managed to help me regain confidence. I am fortunate to have worked by your side: you taught me how to do research thoroughly but also showed me what it meant to persevere in this field.

Thanks to friends and collaborators, I will remember summer 2015 in Sudbury for the wonderful scientists I met, weekends of wildlife watch, and ice-creams by the lakes, rather than for the sweat that went into creating hundreds of mesocosms! In particular, thank you Erik Emilson for showing me how to set up experiments so ingeniously and for being a great friend both in the field and in the office. I also thank my collaborators from Laurentian University: John Gunn, Nate Basiliko, and Nadia Mykytczuk. My Canadian souvenirs would not have been as jolly without your help and without Beth Smith, Kurt Yakimovich, Cyndy Desjardins, and Caroline Emilson.

I wish to thank my wonderful labmates: I feel incredibly lucky to have been surrounded by such clever, funny, and kind people throughout my PhD. In no particular order, thank you Matheus Nunes, Boris Bongalov, Lauranne Gateau Rey, Tom Swinfield, Tommaso Jucker, Carol Ibe, Amelia Fitch, and all the other members and visitors of the Ecology groups, past and present. Javi Igea, you deserve special thanks for your patience with all my questions – from coding to career decisions (and movie recommendations!), you were a genuine help throughout this process. Thanks to my more senior mentors from the Department of Plant Sciences too, David Coomes, Ed Tanner, and Alison Smith – your wisdom was a source of inspiration throughout.

Starting molecular work in an ecology group from scratch would have been virtually impossible without the help of Pawel Baster, Betty Chung, Chris Lambing, Kanu Patel, and Zhengming Wang in the David Baulcombe and Ian Henderson labs, of Moritz Meyer in the Howard Griffiths lab, and of Merlin Sheldrake, Shilo Dickens, and Poppy Marriott for sequencing work. I am also grateful for the NERC-NBAF grant which allowed me to carry out experiments at the University of Sheffield – thank you Helen Hipperson and Gav Horsburgh for welcoming me into the lab, Kirstine Szifris and Lauren Ross for being such great hosts, and massive thanks to Paula Vargas-Pellicer, Päivi Pirhonen, and Yi Zhang for providing the most useful pairs of hands and great company when I had to repeat experiments.

This PhD would never have seen the day without the financial support of my generous sponsors. I thank NERC, the Fondation Marcel Bleustein-Blanchet pour la Vocation, particularly SAS Prince Albert II de Monaco, the Sir Richard Stapley Educational Trust, the Cambridge Philosophical Society, and the British Council for the Entente Cordiale Scholarship they awarded me. I am also immensely appreciative of the Department's help throughout my time there, both financially and administratively. Last but certainly not least, I thank my college, Clare Hall, which not only gave me continuous financial support but also gave me a warm and welcoming environment to live and socialize in throughout my time in Cambridge.

Thank you Alex Assaly. Without you, these last few months would have been much less enjoyable. I will recall the end of my PhD by your side fondly – evenings on the couch writing together next to Sophie. Your support has been invaluable.

Finally, I thank my family for telling me since I was a child that I could do anything I set my mind to and for giving me the opportunities to achieve my dreams. Dad, I am grateful for all that you taught me (jokes included) and for transmitting your

passion for science to me. Mom, you are a model of strength and perseverance who continuously inspired me to overcome the challenges of the thesis. Jonathan, you knew how to comfort me during the tougher periods of this journey. Your joyful wedding was a breath of fresh air in the midst of finishing this thesis and after the difficult months our family went through. I wish you all the best as you embark on your own PhD adventure.

Summary

Terrestrial inputs of organic matter contribute greatly to the functioning of aquatic ecosystems, subsidizing between 30-70% of secondary production. This contribution of terrestrial resources is especially important in boreal lakes that are largely nutrient-poor and thus more responsive to these additions. Yet the mechanisms underlying initial processing of terrestrial resources by microbial communities at the base of lake food webs remain poorly understood. With this in mind, this thesis aims to advance our understanding of lake sediment microbial community assembly and functioning along abiotic gradients, primarily reflecting variation in terrestrial organic matter inputs that are predicted to increase with future environmental change.

Chapter 1 reviews current knowledge on the terrestrial support of lake food webs and highlights gaps in understanding the factors influencing the microbial processing of terrestrial resources. It also provides an overview of metagenomics methods for microbial community analysis and their development over the course of the thesis. **Chapter 2** tests how much of ecosystem functioning is explained by microbial community structure relative to other ecosystem properties such as the present-day and past environment. Theory predicts that ecosystem functioning, here measured as CO_2 production, should increase with diversity, but the individual and interactive effects of other ecosystem properties on ecosystem functioning remain unresolved. **Chapter 3** further questions the importance of microbial diversity for ecosystem functioning by asking whether more diverse microbial communities stabilize ubiquitous functions like CO_2 production and microbial abundances through time. It also aims to identify the biotic and abiotic mechanisms underlying positive diversity-stability relationships. **Chapter 4** then explores how microbial communities assemble and colonize sediments

with varying types and amounts of terrestrial organic matter in three different lakes over a two-month period. Understanding how microbial communities change in relation to sediment and lake conditions can help predict downstream ecosystem functions. Finally, **Chapter 5** discusses the main findings of the thesis and ends with proposed avenues for future research.

Contents

Acknowledgments i

Summary v

Contents vii

List of tables xiii

List of figures xv

Preface xvii

Chapter 1 | General introduction 1

1.1 Allochthony at the center of freshwater ecosystem dynamics 1

1.1.1 Terrestrial resources fuel freshwater ecosystems 1

1.1.2 The movement of allochthonous resources through freshwater food webs . 2

1.1.3 Consequences and limitations of allochthony in lake food webs 4

1.2 Lake sediments as sinks and sources of carbon 5

1.2.1 The contribution of allochthonous resources to carbon cycling 5

1.2.2 The role of microbial communities in mineralization 6

1.3 Lake microbiomes in a changing world 8

1.3.1 Lake sediments in a changing world 8

1.3.2 Biodiversity and ecosystem functioning along environmental gradients 8

1.3.3 The stabilizing effect of diversity on ecosystem functioning 9

1.3.4 Environmental influences on temporal changes in community composition 11

1.4 Advances in studying microbial communities 13

1.5 Thesis aims 17

Chapter 2 | Microbiome functioning depends on individual and interactive effects of the environment and community structure 19

2.1 Abstract..19

2.2 Introduction..20

2.3 Materials and methods...23

2.3.1 Study site 23

2.3.2 Ecosystem functioning and environmental characterization 24

2.3.3 Microbial communities 25

2.3.4 Is there evidence of legacy and environmental effects on community composition? 27

2.3.5 What is the relative importance of different properties for ecosystem functioning? 29

2.4 Results..32

2.4.1 Bacterial community composition 32

2.4.2 Biogeographical patterning of the study sites 33

2.4.3 Linking biogeography and community composition to ecosystem functioning 35

2.4.4 Indirect effects of historical legacies and present-day environment on ecosystem functioning..38

2.5 Discussion...39

Chapter 3 | Biotic and abiotic mechanisms independently stabilize different measures of ecosystem functioning 45

3.1 Abstract..45

3.2 Introduction..46

3.3 Materials and methods...48

3.3.1 Study design 48

3.3.2 Stability and diversity measures 50

3.3.3 Does diversity stabilize ecosystem function as the environment changes? 52

3.3.4 Which proximal mechanisms stabilize ecosystem function? 52

3.4 Results 53

3.4.1 Diversity stabilizes ecosystem function regardless of the environment 53

3.4.2 Undervarying and asynchrony stabilize ecosystem function 54

3.5 Discussion 56

Chapter 4 | Think global, act local: small-scale environmental filters primarily influence microbial community development 61

4.1 Abstract 61

4.2 Introduction 62

4.3 Materials and methods 65

4.3.1 Study design 65

4.3.2 Microbial community characterization 67

4.3.3 Do microbial communities diverge faster from starting leaf material because of differences in sediment or lake conditions? 68

4.3.4 Do microbial communities become dissimilar faster because of differences in sediment or lake conditions? 70

4.3.5 Do microbial taxa form different co-occurrence networks with time because of sediment or lake conditions? 71

4.4 Results 72

4.4.1 Drivers of divergence from original community 72

4.4.2 Community similarity across sediment conditions and lakes 74

4.4.3 Drivers of connectivity among taxa 76

4.5 Discussion 77

Chapter 5 | General discussion 83

5.1 Past historical events directly influence community structure and small-scale conditions drive community changes through time 84

5.2 t-OM quantity and quality are the primary spatial influences over microbial community structure 86

5.3 The environment and community structure influence ecosystem functioning interactively but stabilize it independently 88

5.4 Limitations and avenues for future work 89

5.4.1 Improving experimental design and sequencing 89

5.4.2 Increasing reproducibility with bioinformatics pipelines 91

5.4.3 Advancing research on microbiome functioning 93

5.5 Concluding remarks 94

Appendix A | Biodiversity-ecosystem functioning relationships in lake sediment microbiomes 97

A.1 Supplementary methods 97

A.1.1 Site details..97

A.1.2 Measurement of sediment OM quality 98

A.1.3 Estimation of isotopic mixing model...99

A.1.4 Selection of environmental variables 100

A.1.5 DNA extraction, library preparation, and sequencing 101

A.1.6 Bioinformatics analysis 103

A.2 Supplementary results 111

Appendix B | Diversity-stability relationships in lake sediment microbiomes..121

B.1 Supplementary methods 121

B.1.1 DNA extraction, library preparation, and sequencing 121

B.1.2 Bioinformatics analysis...122

B.2 Supplementary results 123

Appendix C | Changes in lake sediment microbiomes in their early development 125

C.1 Supplementary methods 125

C.1.1 DNA extraction, library preparation, and sequencing 125

C.1.2 Bioinformatics analysis 127

C.2 Supplementary results 128

References 135

List of tables

Table A.1 Among-site variation in 20 environmental variables 105

Table A.2 Correlation matrix for variables displaying among-site variation 107

Table A.3 Environmental gradients across our study lake 107

Table A.4 Gene Ontology categories associated with each of our four subsets of functional genes 108

Table A.5 Model selection statistics and parameter estimates for models of CO_2 production with each of four functional gene subsets 109

Table A.6 Sequencing and taxonomic output of both the shotgun and the amplicon sequencing datasets 111

Table A.7 Community structure across our study lake 112

Table A.8 Partitioning the effect of the past and present environment on bacterial community composition 113

Table A.9 Model averaged parameter effects used to predict CO_2 production....115

Table A.10 Models testing each of the four measures of community structure as a response of the past and present-day environmental predictors 116

Table A.11 Model testing the effect of CO_2 production on the four measures of community structure..118

Table B.1 Model testing ultimate and proximal drivers of CO_2 and community stability 123

Table B.2 Model testing the effects of diversity and the environment on asynchrony 124

Table C.1 Model testing rates of compositional change from the original leaf material per day 129

Table C.2 Associations between environmental conditions and compositional shifts 129

Table C.3 ANOVAs comparing full model to reduced models 130

Table C.4 Model predicting changes in community dissimilarity 131

Table C.5 Model predicting each kingdom's abundance with time 132

Table C.6 Model testing the effects of t-OM quantity, t-OM quality, and lake, respectively, on changes in degrees with time 133

Table C.7 Taxonomic identification of the 10 most abundant ASVs 134

List of figures

Figure 1.1 Aquatic food web with allochthonous inputs 4

Figure 1.2 Pathways influencing community structure and ecosystem function ...13

Figure 1.3 Innovations in microbial community sequencing 16

Figure 2.1 Conceptual model of pathways by which different ecosystem properties affect ecosystem functioning 32

Figure 2.2 Bacterial communities across geographic and environmental space 34

Figure 2.3 Bacterial communities across sites and environments 35

Figure 2.4 Effects of the environment and microbial communities on ecosystem functioning 37

Figure 2.5 Drivers of ecosystem functioning in littoral lake sediments 38

Figure 3.1 Stability of CO_2 production as a function of mean OTU richness 53

Figure 3.2 The mechanism underlying the stabilizing effect of diversity on CO_2 production..54

Figure 3.3 Relationship between community stability and its standard deviation as a function of asynchrony..55

Figure 3.4 Pathways by which diversity directly stabilizes CO_2 production and t-OM quantity indirectly stabilizes microbial community abundance ..56

Figure 4.1 Microbial community divergence from the original leaf community ...73

Figure 4.2 Microbial community dissimilarity with time at different levels of t-OM quantity and quality 75

Figure 4.3 Microbial community dissimilarity with time in different lakes...........76

Figure 4.4 Microbial co-occurrence patterns with time...77

Figure A.1 Map of the study sites..97

Figure A.2 Frequency distribution of family-level presence across sites..............112

Figure C.1 Sediment pore water conditions in the three lakes..............................128

Preface

To facilitate publication of the thesis, **Chapters 2-4** are written as manuscripts for peer-reviewed journals. Co-author contributions included field assistance, laboratory support, data provision, and supervisory guidance. As a result, I use the pronoun "we" rather than "I" in these chapters. Specifically, I collected the data for **Chapters 3** and **4**, extracted the DNA and prepared sequencing libraries for **Chapters 2-4**, sequenced the shotgun libraries for **Chapters 2** and **3**, and carried out all bioinformatics, statistical analyses, and writing for **Chapters 2-4** with input from Andrew Tanentzap.

At the time of submission, **Chapter 2** has been published:

Orland C, Szkokan-Emilson EJS, Basiliko N, Mykytczuk N, Gunn J, Tanentzap AJ Microbiome functioning depends on individual and interactive effects of the environment and community structure. *ISME J* 2018, in press, https://doi.org/10.1038/s41396-018-0230-x

Designed the study: CO with support from EJSES, NB, NM, JG, and AJT.

Collected the data: EJSES and AJT. *Analysed the data:* CO with support from EJSES, NB, JG, and AJT. *Wrote the paper:* CO with support from EJSES and AJT.

In addition, the experimental procedure followed in **Chapters 3** and **4** is largely based on a paper to which I contributed as a co-author:

Tanentzap AJ, Szkokan-Emilson EJ, Desjardins CM, **Orland C**, Yakimovich K, Dirszowsky R, Mykytczuk N, Basiliko N, Gunn J. Bridging between litterbags and whole-ecosystem experiments: A new approach for studying lake sediments. *J Limnol* 2017; **76**: 431–437.

Lastly, the following papers have resulted from work conducted during the course of the thesis but are not included in the text:

Fitch A, **Orland C**, Willer D, Emilson EJS, Tanentzap AJ. Feasting on terrestrial organic matter: dining in a dark lake changes microbial decomposition. *Glob Chang Biol* 2018; **24**: 5110—5122.

Jucker T, Wintle B, (…), **Orland C**, (…), Mukherjee N. Ten-year assessment of the 100 priority questions for global biodiversity conservation. *Cons Biol* 2018; in press, https://doi.org/10.1111/cobi.13159

Jones JA, Wei XA, (…), **Orland C**, (…), Archer E. Climate-Forest-Water-People Relations: Seven System Delineations. 2018; In *Forest and Water on a Changing Planet: Vulnerability, Adaptation and Governance Opportunities*, p.27.

Blanco JA, Bishop K, (…), **Orland C**, (…), Li Q. Forest Landscape Hydrology in a 'New Normal' Era of Climate and Land Use Change. 2018; In *Forest and Water on a Changing Planet: Vulnerability, Adaptation and Governance Opportunities*, p.81.

Antwis RE, Griffiths SM, (…), **Orland C**, (…), Sutherland WJ. Fifty important research questions in microbial ecology. *FEMS Microbiol Ecol* 2017; **93**: 1-10.

Chapter 1

General introduction

1.1 Allochthony at the center of freshwater ecosystem dynamics

1.1.1 Terrestrial resources fuel freshwater ecosystems

Organic matter derived from terrestrial vegetation and soil, such as wood debris and animal detritus, subsidizes much of the secondary production in aquatic food webs. Two lines of evidence support this observation. First, in half of all cases, an average of 42% of the biomass of lake zooplankton can be traced to allochthonous resources exported from terrestrial ecosystems as opposed to the autochthonous resources produced within the boundaries of freshwater ecosystems (Tanentzap et al. 2017). Multiple isotopic analyses also revealed that fish biomass was derived from terrestrial resources at the stream-lake interface (e.g. 57% of fish biomass for Karlsson et al. 2012, 34-66 % for Tanentzap et al. 2014). Second, many studies provide evidence that respiration can exceed primary production in freshwater systems, particularly in lakes with low primary productivity (Cole et al. 1994, del Giorgio et al. 1997, Jansson et al. 2000, Duarte and Prairie 2005, Tranvik et al. 2009). As more carbon was released than could be predicted from the biomass of photosynthetic organisms within the lake, respiration must have resulted from decomposition of terrestrial organic matter (t-OM) inputs that constitute an organic carbon (C) source.

The reliance on allochthonous resources is especially strong in northern temperate regions like the boreal (latitude 50° to 70°N), where lakes tend to be less productive (Jonsson et al. 2001, Einola et al. 2011) and receive large amounts of t-OM

relative to the area they occupy (Wurtsbaugh et al. 2002). In boreal lake food webs, allochthonous subsidies from riparian litterfall account for up to half of the resources available to aquatic organisms (Karlsson et al. 2012). Understanding these terrestrial-aquatic linkages is therefore essential for predicting how future environmental changes to terrestrial catchments will affect boreal lakes, which are ecologically and economically important as they store more than 60% of the planet's freshwater (Schindler and Lee 2010).

1.1.2 The movement of allochthonous resources through freshwater food webs

Allochthonous inputs tend to be exported into lakes from drainage streams rather than surface runoff (Dillon and Molot 1997). As streams feed into lakes, reductions in flow rates result in large amounts of these allochthonous inputs settling near shore, creating delta landforms (Szkokan-Emilson et al. 2011). The littoral (or nearshore) zone of lakes thus disproportionately accumulates allochthonous inputs, and littoral sediment constitutes the site where most terrestrially-derived material will be transferred into the food web (Wetzel 2001). Importantly, up to half of the particulate organic carbon (POC) in lakes originates from terrestrial sources (Pace et al. 2004). POC will be broken down into smaller forms differently according to its structure (e.g. particle size, material), mixing and residence time (Amon and Benner 1996), and these characteristics strongly depend on whether it is from aquatic or terrestrial origin (Wetzel 1995). The microbial species present in the sediment will also influence the degradation of POC (Williams et al. 2010).

Additionally, inland waters receive inputs of terrestrially dissolved organic carbon (DOC), which increase with vegetation cover, conductivity, soil C:N ratio, soil carbon, and soil wetness (Sobek et al. 2007, Tanentzap et al. 2017). Degradation of within-lake POC also contributes to the DOC pool (Pace et al. 2004). DOC is a

component of dissolved organic matter (DOM). DOM is comprised of low molecular weight compounds, such as amino acids and carbohydrates, and high molecular weight compounds, such as lignin and aromatic hydrocarbons (Evans et al. 2005). The structure of DOM varies greatly with the origin of the DOM and determines the amount of DOC available and how it will be processed by microbes (Kellerman et al. 2015), subsequently affecting the rest of the food web.

POC and DOC can enter aquatic food web in two main ways (Fig. 1.1). First, by degrading organic C, microbes produce new biomass that mobilizes terrestrial resources into the aquatic food web (Hessen 1998, Jansson et al. 2007). In the process, microbes release inorganic nutrients that are made accessible to algae, for example by respiring inorganic forms of C, such as CO_2 (Cole et al. 1994, Roehm et al. 2009). Dissolved inorganic carbon (DIC) can indeed fuel the fixation of new sources of organic carbon by algal photosynthesis (Hanson et al. 2006). Second, organic C can be directly ingested by macroinvertebrates and zooplanktonic primary consumers (Cole et al. 2006). Although the incorporation of t-OM into aquatic food webs is well understood, the mechanisms of this energy transfer, e.g. whether it depends on abiotic or biotic conditions, remains poorly studied.

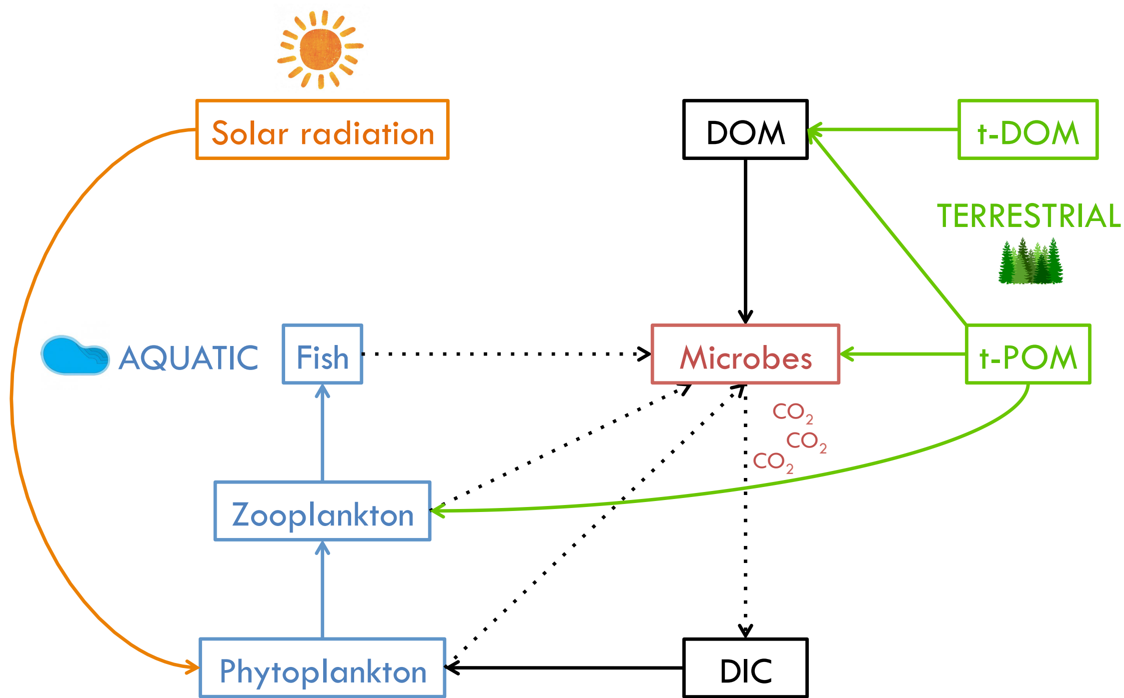

Fig. 1.1 | Aquatic food web with allochthonous inputs (t-DOC and t-POC) and energy moving upwards through the food web. Solid arrows represent the pathway of entry of carbon to higher trophic levels and dashed arrows represent bacterial activity that "loops" the energy from detritus back into the food web. Allochthonous sources are directly mobilized into aquatic food webs by microbes and invertebrates, thereby tying land to water.

1.1.3 Consequences and limitations of allochthony in lake food webs

Evidence that t-OM subsidizes higher trophic levels in the freshwater food web suggests that allochthonous resources must be entering lakes at the base of the food web (Richardson 1991, Carpenter et al. 2005). For example, Tanentzap et al. (2014) found that the sizes of juvenile fish increased as more t-OM was exported into the nearshore waters beneath eight boreal catchments, with energy flowing upwards through bacteria and zooplankton communities. These results supported the trophic upsurge hypothesis (Baranov 1961), whereby differences in the productivity of biological communities are due to differing amounts of energy transferred upwards through the food web rather than to top-down changes exerted by predators. Microbial communities, situated at the

bottom of the food web, should therefore play a key role in influencing entire food webs and ecosystems more broadly.

There may also be limits to whether OM quantity benefits aquatic food webs. Terrestrial inputs that are too large can shade phytoplankton and limit aquatic production (Karlsson et al. 2009, Jones et al. 2012). In contrast to the Tanentzap et al. (2014) study, Finstad et al. (2014) found that once forest cover reached a certain threshold density, it limited fish biomass rather than enhanced it due to light limitation of primary production at increasing depth. Additionally, reduced water clarity may shift planktonic species to less nutritious ones with lower fatty acid concentrations, which are essential to support higher trophic levels (Brett and Muller-Navarra 1997). Zooplankton preferentially feed on phytoplankton, meaning that allochthonous inputs, which are of much lower nutritional quality, may be less selected if phytoplankton are available (Brett et al. 2009, Brett et al. 2012). Abiotic factors that are influenced by t-OM like temperature and UV levels may also affect primary production within the system and subsequently higher trophic levels (Lefébure et al. 2013). Whether this allochthonous organic carbon is adding to autotrophic fixation by algae, replacing it, or perhaps even limiting it (Finstad et al. 2014), remains uncertain, and may depend on the role of microbial communities. Most studies have focused on the positive relationship between terrestrial OM inputs and secondary consumer abundance, however the role that microbial organisms play in enabling this transfer of energy and the link between OM and microbes needs to be better understood.

1.2 Lake sediments as sinks and sources of carbon

1.2.1 The contribution of allochthonous resources to carbon cycling

In addition to supporting lake food webs, microbial decomposition of POM and DOM in sediments contributes to whole-lake nutrient and carbon cycling. The surface area of

lakes makes up 0.6% of the oceans' surface area (Einsele et al. 2001). Yet lakes store disproportionally high quantities of carbon due to the rapid accumulation of sediments and high preservation times (about 10,000 years), burying the equivalent of a fourth of the annual atmospheric carbon stored in oceans (Dean and Gorham 1998, Einsele et al. 2001). Of these carbon inputs, lakes have been estimated to receive 5.1 $Pg.C.year^{-1}$ from terrestrial sources, with approximately 0.6 of which is buried in the sediment (Battin et al. 2009, Tranvik et al. 2009, Drake et al. 2017). In boreal lakes, 40-70% of the carbon buried in the sediments will be respired, a figure which decreases with within-lake productivity (Pace and Prairie 2005).

Lakes play two contrasting roles in carbon cycling: they fix carbon through photosynthesis, but also mineralize organic carbon, releasing DIC (e.g. CO_2, CH_4) through respiration. The balance between fixation and mineralization establishes whether there is a net uptake or net loss of CO_2. Yet accurately estimating this balance is challenging, as many components (e.g. allochthonous sources, losses from flow, sediment burial) are difficult to measure. Hanson et al. (2015) identified 10 questions that need to be addressed to calculate lake carbon budgets more accurately. Amongst these, quantifying the partitioning of allochthonous inputs to lakes and understanding the observed variation in gas exchange through mineralization at different sites are key questions to answer. These processes are heavily reliant on the gas fluxes from sediment microbial communities, and any changes in decomposition rates will consequently impact biogeochemical cycles.

1.2.2 The role of microbial communities in mineralization

Microbial communities drive decomposition in sediment, and different taxa contribute differently to this process. Heterotrophic bacterial groups, which consume a large portion of the DOM in the sediment (Solomon et al. 2015), have been recognized as

active decomposers of low molecular weight compounds (Berggren et al. 2010a) based on nutrient requirements (Peura et al. 2012) and on their metabolic activity (Hooper et al. 2002). For example, *Betaproteobacteria* and *Bacteroidetes* are found in aquatic sediments rich in organic carbon and the *Nitroso*-genera is known to oxidise ammonia (Kirchman 2002, Fazi et al. 2005, Fierer et al. 2007). Microcosm and field experiments in streams have also shown that fungi convert 14-48% of organic carbon from leaves into CO_2 (Suberkropp 1991, Gulis and Suberkropp 2003). Additionally, aquatic fungal species actively uptake nitrogen (Tank and Dodds 2003), and this nitrogen mobilization is itself related to higher levels of respiration (Stelzer et al. 2003) and leaf litter decomposition (Huryn et al. 2002), contributing to OM decomposition as part of the microbial loop (Pomeroy 1974, Azam et al. 1983, Baldy et al. 2002, Wurzbacher et al. 2010). Finally, archaeal taxa, particularly methanogens that metabolize carbon into methane, are responsible for 10-50% of the overall carbon mineralization in lakes (Bastviken et al. 2008).

The range of different functions (i.e. functional diversity) performed within microbial communities will likely influence mechanisms of decomposition. For example, sulfate-reducing bacteria, which can decompose OM anaerobically, are highly important in freshwater sediments, which are usually rich in sulfate. These bacteria have been shown to compete with methanogens for hydrogen, which may have consequences for oxidation processes (Lovley and Klug 1983). The importance of methane-oxidizing bacteria has also been highlighted in lakes, with methanogenic carbon shown to support a substantial part of production in higher trophic levels (17% and 12%, Ravinet et al. 2010 and Lau et al. 2014, respectively, and up to 46% in the summer in chalk rivers as shown by Shelley et al. 2014). Overall, the variety of functions performed by microbial communities in freshwater sediments suggests that specific microbial species may be specialized, or more frequently associated, with certain tasks. If these different tasks are

decomposition-related, decomposition may be greater in species-rich communities because of the complementarity of functions conferred by different species.

1.3 Lake microbiomes in a changing environment

1.3.1 Lake sediments in a changing world

Boreal ecosystems are undergoing rapid changes (Kirtman et al 2013), with longer growing seasons and shifts in tree species composition from coniferous to deciduous (Heathcote et al. 2015). Consequently, the type and amount of allochthonous resources delivered to inland waters and its impact on terrestrial-aquatic linkages is expected to change (Boisvert-Marsh et al. 2014, Creed et al. 2018). Other abiotic processes – mainly due to anthropogenic activities – are expected to impact boreal lake ecosystems. For example, warming and increased reactive nitrogen deposition are accelerating carbon burial in sediments (Heathcote et al. 2015). Additionally, increased exports of DOC due to, among other factors, precipitation-driven runoff and reductions in atmospheric acid deposition, are responsible for increased "browning" of lakes (Clark et al. 2010, Finstad et al. 2016, Creed et al. 2018). All these environmental changes are likely to change the biodiversity, functioning, and stability of lake sediments.

1.3.2 Biodiversity and ecosystem functioning along environmental gradients

Decomposition rates depend on the presence of specific microbial groups (Hooper et al. 2002, Krause et al. 2014), but whether these rates increase systematically with the diversity and abundance of microbial taxa and genes (hereafter "community structure" as defined by Bier et al. 2015) is poorly understood. Diverse microbial communities may elevate decomposition rates if they capture a greater range of species capable of breaking down carbon substrates. Such a positive association between biodiversity and ecosystem functioning (B-EF) may arise through several mechanisms. For example,

with more species present, the community may be dominated by species with traits that enhance productivity (Loreau and Hector 2001, Cardinale et al. 2002). Species can also coexist without competing, by using resources in a complementary way, through ecological niche partitioning and facilitation (Loreau and Hector 2001). On the other hand, direct competition for resources and overlapping niche space with increased diversity can lead to functional redundancy and competitive exclusion and result in a negative B-EF relationship (Naeem 2002).

Importantly, B-EF relationships can shift along environmental gradients (Hooper and Dukes 2004, Graham et al. 2016; Fig. 1.2). In such cases, species richness may not be the best predictor of ecosystem function and species composition or species traits may be better (Purvis and Hector 2000, Stachowicz et al. 2007, Solan et al. 2013). Species richness can also change along resource gradients. For example, the stress-gradient hypothesis (SGH) predicts that biodiversity should be higher in resource-limited environments due to facilitation and lower in less-stressful environments where competitive interactions are frequent (Bertness and Callaway 1994). Any such change to biodiversity along environmental gradients may subsequently influence ecosystem functioning (Callaway 2007, Maestre et al. 2009, Jucker and Coomes 2012).

1.3.3 The stabilizing effect of diversity on ecosystem functioning

A key finding from research into B-EF is that, besides promoting ecosystem functioning, the presence of more species stabilizes ecosystem function through time (Tilman 1999, Isbell et al. 2009, Hautier et al. 2014). Most of the evidence for the stabilizing effect of diversity on ecosystem function comes from grassland studies, where species-rich communities fluctuate less in their primary productivity over time than species-poor ones (Hautier et al. 2014, Jucker et al. 2014). However, much less is known about the diversity-stability (D-S) relationship in microbes (Downing et al.

2014, Wagg et al. 2018). With environmental changes expected to drastically alter species diversity in the future (Cardinale et al. 2012), understanding the D-S relationship and its mechanisms will improve predictions of ecosystem processes carried out by microbial communities (McGrady-Steed et al. 1997, McCann 2000).

More diverse microbial communities can stabilize ecosystem functions like decomposition through at least three mechanisms that have been identified from other systems (Downing et al. 2014). First, ecosystem function may be stabilized by diversity if the presence of more species increases mean ecosystem function over time (Hector et al. 2010). Similar to the B-EF relationship, this stabilizing effect may arise if species partition resources such that they minimize inter-specific competition and have greater than expected functioning (Loreau and Hector 2001). Second, increased diversity may reduce the variability of ecosystem function by enhancing facilitative interactions among species and thus promoting species coexistence (Mulder et al. 2001, del Río et al. 2014). Finally, more diverse communities can maintain ecosystem function because there are more species that can increase in abundance and compensate for declines in the abundance of other species that contribute to functioning (Houlahan et al. 2007, Hector et al. 2010).

Environmental change may also influence the mechanisms underlying the stabilizing effect of diversity (McGrady-Steed et al. 1997, McCann 2000; Fig. 1.2). One reason why these diversity-stability (D-S) relationships might vary in boreal freshwaters is because larger inputs of aromatic t-OM are forecast (Creed et al. 2018). As t-OM has been shown to be bioavailable for microbial growth and respiration (Guillemette and del Giorgio 2011, Lapierre et al. 2013), such environmental changes may strengthen the stabilizing effect of diversity by providing more resources that will enhance mean ecosystem function (Micheli et al. 1999, Downing et al. 2008). D-S relationships can also be strengthened by environmental changes that promote species coexistence. Such

circumstances might arise if a species mitigates the negative effects of an environmental change on the abundance of other species, subsequently reducing temporal variability in ecosystem function (Vogt et al. 2006, Romanuk et al. 2010, Downing et al. 2014). Finally, environmental changes that differentially modify species abundances and subsequently enhance species fluctuations can also influence the strength of D-S relationships (Ives et al. 1999, Thébault and Loreau 2005, Loreau and de Mazancourt 2008). Despite the importance of predicting ecosystem functioning in the context of a changing world, few studies have simultaneously tested for these mechanisms and their environmental dependency outside of primary producers (Grman et al. 2010, Downing et al. 2014).

1.3.4 Environmental influences on temporal changes in community composition

How communities change with time will subsequently shape the traits and functions that they perform (Fukami and Morin 2003, Fukami et al. 2010, Popp et al. 2017), influencing both community structure and ecosystem function. The temporal processes that influence how these communities change, like turnover and succession rates, are expected to be high in microbial communities (Schmidt et al. 2007). Additionally, small- and large-scale abiotic conditions that filter species assemblages will influence these temporal changes in community composition (Langenheder and Székely 2011, Kraft et al. 2015). For example, large-scale filters, such as connectivity among catchments and lake conditions (Nelson et al. 2009, Nino-Garcia et al. 2016), and smaller-scale filters, such as resource quantity and quality (Ruiz-Gonzalez et al. 2015a, Tanentzap et al. 2014) and pH (Fierer and Jackson 2006), structure microbial communities. Yet little is know about how these different environmental scales influence the rate at which communities develop with time.

Communities can also change temporally when the presence of one species affects the presence of others (Newman 2003, Gotelli and McCabe 2002). Priority effects, whereby the presence of one species in a habitat reduces the probability that another species will colonize that habitat, explain how taxa that first colonize a habitat and start occupying niche space may be more likely to dominate (Shulman et al. 1983). These priority effects mean that the early stages of development are particularly important for later community composition (Fierer et al. 2010, Fukami 2015). Additionally, positive interactions among species may intensify through time, as functional complementarity among coexisting species strengthens and redundancy diminishes (Gross et al. 2014, Zuppinger-Dingley et al. 2014). These interactions can generate co-occurrence relationships that will in turn impact aspects of the community such as taxonomic and functional composition (Chesson 2000, Williams et al. 2014) and, importantly, alter the stability of the community over long timescales (HilleRisLambers et al. 2012). Environmental filters will also influence these inter-specific interactions by indirectly enhancing or limiting resource availability and thus modifying which species will persist in the community (Tilman 1982, HilleRisLambers et al. 2012). Consequently, understanding how microbial communities change in their early development along environmental gradients will improve predictions of both community structure and ecosystem functioning (Fig. 1.2). These predictions will be particularly valuable where small- and large-scale changes are forecast, such as in boreal ecosystems where warmer climates and increased t-OM export will impact freshwaters (Creed et al. 2018).

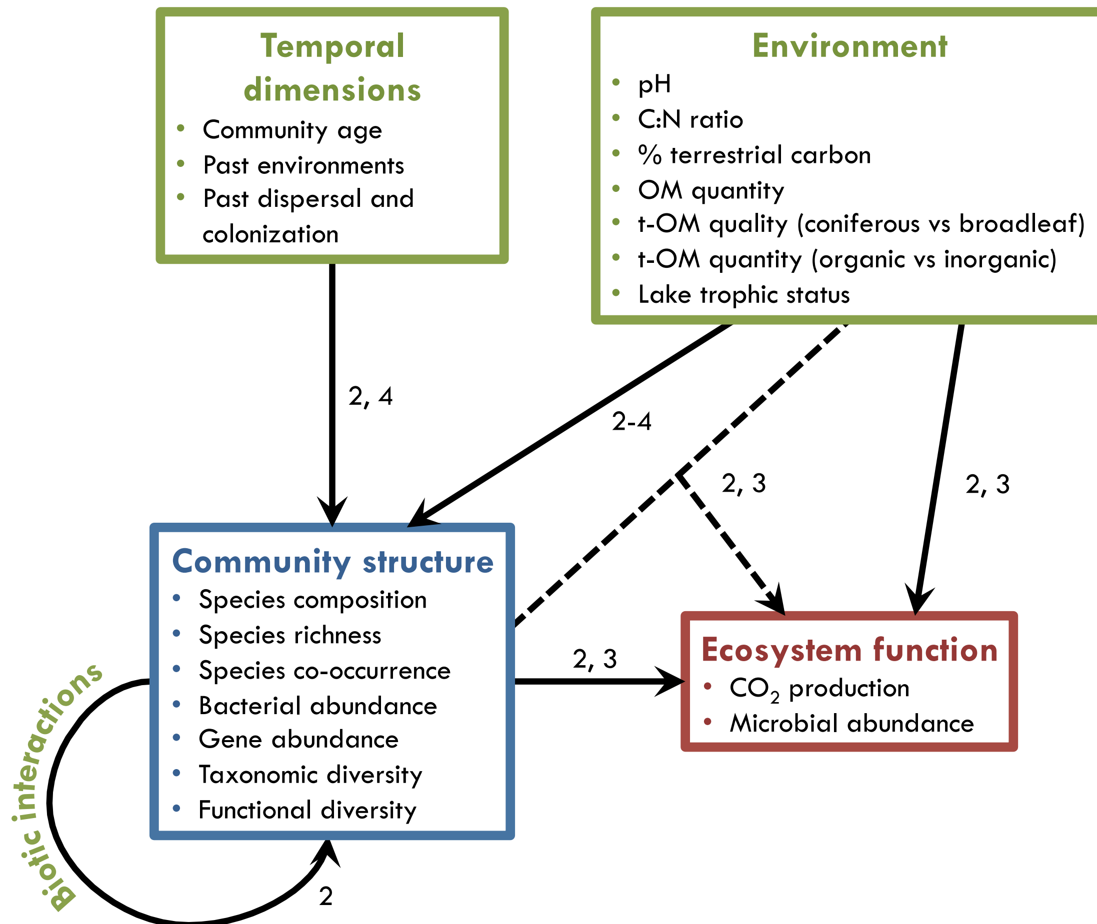

Fig. 1.2 | Pathways by which the environment and temporal dimensions affect community structure and ecosystem function. Solid arrows represent pathways tested in this thesis (with corresponding chapter numbers) and dashed arrows represent interactive effects. Green boxes and text indicate processes informing community structure, and the blue box and red box indicate components of community structure and ecosystem function, respectively.

1.4 Advances in studying microbial communities

Sediment microbes are essential to understanding terrestrial-aquatic linkages as they are the primary organisms that interact with t-OM. Traditionally, it has been difficult to gather information on microbial communities. Identifying them taxonomically was complicated by their microscopic size and relative morphological similarity, and identifying them functionally was limited by the difficulties associated with microbial culturing. Over the last three decades, advances in DNA sequencing technologies and

bioinformatics have however transformed the field of microbial ecology. The composition and dynamics of complex communities from a range of environments are being studied in greater detail than ever before (Solieri et al. 2013, Knight et al. 2018). In particular, the commercialization of the first high-throughput sequencer in 2005 (i.e. 454 Roche) progressively replaced more traditional molecular methods like terminal restriction fragment length polymorphism (T-RFLP) and denaturing gradient gel electrophoresis (DGGE), and, in doing so, revolutionized our understanding of microbes. Reductions in cost and easy-to-use library preparation kits meant sequencing hundreds of samples to depths of tens of thousands of reads rapidly became feasible both financially and technically.

The shift from fingerprinting methods (e.g. T-RFLP and DGGE), which allowed diversity patterns in communities to be determined but not the identity of the taxa present, to sequencing methods that recovered taxonomic and functional information had a major impact on thinking in ecology. These new sequencing methods immediately advanced our understanding of microbial communities as a whole, providing information on their structure, diversity, and on their rarer components at a much greater depth than previously possible (Pedrós-Alió 2007, del Giorgio 2010). Perhaps one of the first – and certainly one of the most cited – papers to reveal the potential of metagenomics dates back to 2004, when Venter et al. sequenced seawater samples from the Sargasso Sea. They discovered over 1.2 million previously unknown genes and demonstrated that this method provided relatively unbiased taxonomic information, thereby highlighting the importance of metagenomics for our understanding of species and gene diversity in the environment (Venter et al. 2004). Another seminal paper provided the first evidence for global patterns of bacterial diversity in seawater and identified the environmental drivers associated with these patterns (Zinger et al. 2011). In the field of freshwater ecology, some of the first, key pieces of research done with

metagenomics revealed: 1) support for the partitioning of bacterial taxa along carbon substrate resource gradients (Jones et al. 2009), 2) the prevalence of bacterial dormancy in nutrient-poor lakes and the disproportionate activity of rare taxa relative to common bacteria (Jones et al. 2010), and 3) bacterial resistance to whole-ecosystem disturbance (Shade et al. 2012).

More recent developments have led to further improvements. While it was only possible to recover community-wide taxonomic information from sequencing technologies as recently as 2013, it is now possible to obtain high-resolution functional data on microbial communities (Tessler et al. 2017; Fig. 1.3). This shift from "amplicon" to "shotgun" sequencing emerged from innovations in library preparation and sequencing platforms. Amplicon sequencing involves targeting a conserved but highly variable region of a single gene – for example, 16S rRNA and ITS genes are typically used for bacteria and fungi, respectively. By contrast, shotgun sequencing indiscriminately sequences across entire genomes. There are multiple advantages to this. First, shotgun library preparations do not require a PCR amplification step, which can introduce multiple biases (e.g. amplification bias due to primer affinity for certain sequences, amplicon size, and number of PCR cycles, as well as bias induced by differential gene copy numbers among taxa; Clooney et al. 2016, Knight et al. 2018). Second, considering no single gene is targeted, information on thousands of genes can be obtained and thus simultaneously provide both functional and taxonomic information. The downsides to shotgun sequencing (i.e. relatively higher costs and challenges in storing and processing the much larger datasets) mean that amplicon sequencing is still widely used. Yet there appears to be increasing support towards using shotgun sequencing approach (Clooney et al. 2016, Ranjan et al. 2016, Knight et al. 2018). This approach is particularly valuable for research on ecosystem functioning as it

provides information on the microbial genes present that can be linked to processes like decomposition.

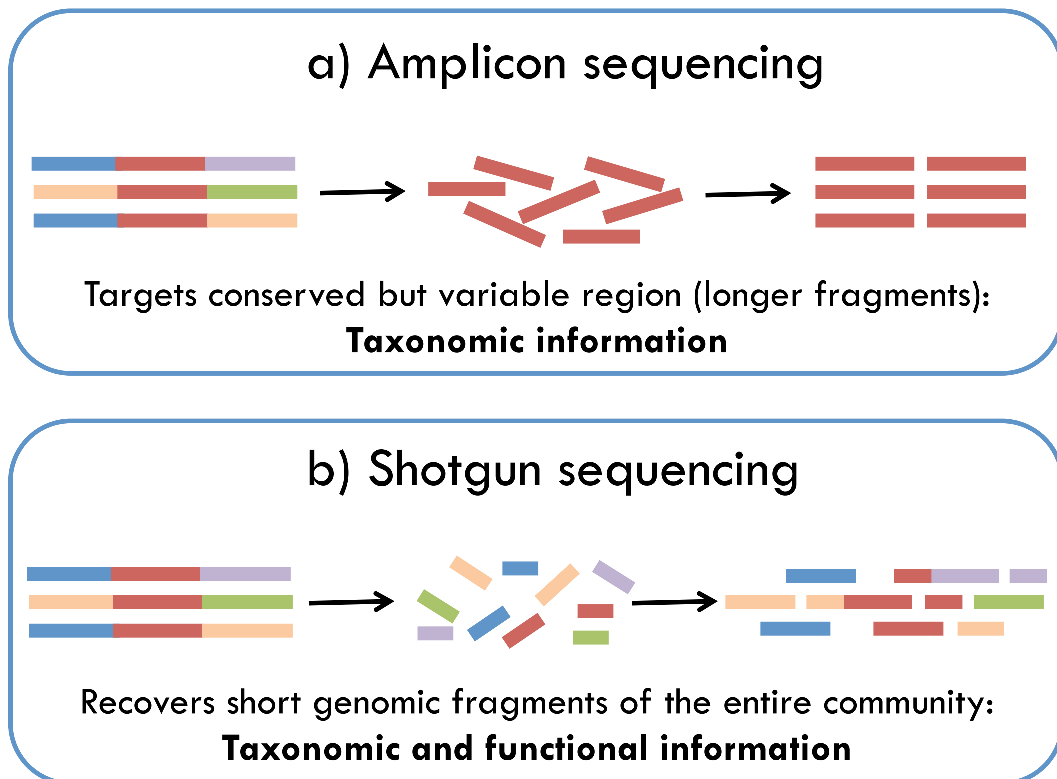

Fig. 1.3 | Innovations in microbial community sequencing are revolutionizing our understanding of their structure and dynamics. While a) amplicon sequencing is a fast, cost-effective, and well-developed method of obtaining taxonomic information, b) shotgun sequencing reduces biases associated with amplification and allows functional information to be simultaneously recovered.

The field of bioinformatics – i.e. the computational processing of sequencing data – has also evolved hugely in recent years. With all its advantages, high-throughput sequencing also brought about a major challenge: recovering taxa and functions from the millions of bases sequenced. Numerous tools and pipelines were developed to make these steps more accessible and computationally efficient. One such novelty worth highlighting is DADA2 (Callahan et al. 2016a), which has already gained in popularity

in the field of microbial ecology (Hugerth and Andersson 2017, Knight et al. 2018, Pollock et al. 2018). Rather than arbitrarily determining taxa with the usual 97% threshold sequence similarity, this method defines unique sequence variants thanks to a method that controls for amplification and sequencing errors (Callahan et al. 2016a). The sequences produced are thus more biologically meaningful as they represent unique microbial taxa, and yield more accurate and reproducible amplicon data across studies than previous methods allowed (Callahan et al. 2017). Obtaining higher resolution taxonomic information will improve estimates of microbial richness and thus our understanding of microbial community structure.

Microbial ecology stands out as a field currently undergoing particularly fast-paced and innovative changes, with more promising technologies on the horizon (e.g. improved methodologies for transcriptomics, low-cost third generation sequencing). These innovations will allow microbes to be sampled at increasingly high-resolution, and to link microbial community structure and dynamics to key ecosystem processes, such as decomposition.

1.5 Thesis aims

The aim of this thesis is to address how microbial communities assemble and function along terrestrial resource gradients in boreal lake sediments. Applying advances in microbial genomics to field data from observational and experimental studies, I address three main questions:

1) How do temporal processes influence aquatic microbial community assembly and structure (**Chapters 2** and **4**)?
2) How do spatial environmental gradients related to terrestrial inputs influence community structure (**Chapters 2-4**)?
3) How do the environment and community structure individually and interactively influence ecosystem function (**Chapters 2** and **3**)?

More specifically, in **Chapter 2** I test how much of ecosystem functioning is explained by microbial community structure relative to other ecosystem properties such as the present-day and past environment. Theory predicts that ecosystem functioning, here measured as CO_2 production, should increase with diversity, but the individual and interactive effects of other ecosystem properties on ecosystem functioning remain unresolved. In **Chapter 3** I further question the importance of microbial diversity for ecosystem functioning by asking whether more diverse microbial communities stabilize important ecosystem functions over time and how this stabilizing effect might vary along environmental gradients. I also aim to identify the biotic and abiotic mechanisms underlying positive diversity-stability relationships. **Chapter 4** then explores how microbial communities colonize sediments with replicated gradients of terrestrial organic matter in three lakes with differing water quality. Understanding how microbial communities change over time in relation to small-scale (i.e. sediment conditions) and large-scale (i.e. lake conditions) environmental filters can help predict downstream ecosystem functions. Finally, I discuss the main findings of the thesis and end with proposed avenues for future research in **Chapter 5**.

Chapter 2

Microbiome functioning depends on individual and interactive effects of the environment and community structure

2.1 Abstract

How ecosystem functioning changes with microbial communities remains an open question in natural ecosystems. Both present-day environmental conditions and historical events, such as past differences in dispersal, can have a greater influence over ecosystem function than the diversity or abundance of both taxa and genes. Here, we estimated how individual and interactive effects of microbial community structure defined by diversity and abundance, present-day environmental conditions, and an indicator of historical legacies influenced ecosystem functioning in lake sediments. We studied sediments because they have strong gradients in all three of these ecosystem properties and deliver important functions worldwide. By characterizing bacterial community composition and functional traits at 8 sites fed by discrete and contrasting catchments, we found that taxonomic diversity and the normalized abundance of oxidase-encoding genes explained as much variation in CO_2 production as present-day gradients of pH and organic matter quantity and quality. Functional gene diversity was not linked to CO_2 production rates. Surprisingly, the effects of taxonomic diversity and normalized oxidase abundance in the model predicting CO_2 production were attributable

to site-level differences in bacterial communities unrelated to the present-day environment, suggesting that colonization history rather than habitat-based filtering indirectly influenced ecosystem functioning. Our findings add to limited evidence that biodiversity and gene abundance explain patterns of microbiome functioning in nature. Yet we highlight among the first time how these relationships depend directly on present-day environmental conditions and indirectly on historical legacies, and so need to be contextualized with these other ecosystem properties.

2.2 Introduction

Biodiversity-ecosystem functioning (B-EF) relationships are generally expected to be positive because more unique functions are captured as species numbers increase. While this prediction often holds true for macroorganisms (Tilman et al. 2014), it is still contested for microorganisms. Some have found support for positive B-EF relationships in microbial communities (Bell et al. 2005, Venail and Vives 2013, Delgado-Baquerizo et al. 2016, Laforest-Lapointe et al. 2017), as expected if microbes perform a diversity of functions (e.g. litter decomposition, temperature regulation, nutrient cycling), which increase with numbers of taxa. However, others have found negative and no B-EF relationships (Jiang 2007, Becker et al. 2012). One explanation for this conflicting evidence is that taxonomic diversity may have relatively little influence on functioning in microbial ecosystems that are saturated by thousands of species that overlap in their traits (Nielsen et al. 2011). Empirical evidence to support B-EF theory in microbes has also come from communities where species richness rarely exceeds 100 taxa (Krause et al. 2014), which is much less than the thousands of taxa found in natural communities, e.g. up to 9,000 prokaryotic taxa in 1 cm^3 of soil (Bardgett and van der Putten 2014), but see Delgado-Baquerizo et al. (2016) and Laforest-Lapointe et al. (2017).

Functional information may help resolve the lack of widespread evidence for a positive B-EF relationship in microbial communities by overcoming the limitations of using solely taxonomic information. In particular, taxonomic information may have little value where it does not map onto function, and this may be relatively common in microorganisms because taxa are delineated from classifying closely-related genetic sequences rather than morphological or physiological traits (Martiny et al. 2013). Horizontal gene transfer can also complicate the use of taxonomic information in microbes because traits might not be vertically transmitted as expected based on phylogeny (Doolittle 1999). Given these concerns, Graham et al. (2016) recently found that combining both taxonomic and functional measures of diversity strengthened predictions of ecosystem functioning across 82 microbial systems compared to models including only microbial biomass. While 56% of the variation in functioning was explained by environmental variables, such as pH and temperature, incorporating information about microbial taxonomic diversity explained, on average, 8% of additional variance (Graham et al. 2016). Total functional gene abundance equally improved predictions of microbial respiration (Graham et al. 2016). These findings underscore the importance of considering abundance *and* diversity metrics of both function *and* taxonomy when predicting B-EF relationships.

Past events can also leave a legacy on present-day microbial communities and influence species composition and subsequent ecosystem function as much as the contemporary environment (Vass and Langenheder 2017, Martiny et al. 2017). These events can include past differences in dispersal and environmental conditions that have differentially sorted species composition. Thus, microbial communities can have less gene flow and greater genetic divergence as they become increasingly distant in space irrespective of environmental similarity (Martiny et al. 2006). For example, recent evidence has shown that historical legacies can result in different microbial

communities in similar environmental conditions, partly because microbes are not ubiquitously distributed (Friedline et al. 2012). Consequently, B-EF relationships have been found to depend on the temporal order of community assembly (Fukami and Morin 2003, Fukami et al. 2010) and to vary over small spatial distances (i.e. 20 m; Lear et al. 2014). Despite their potential importance for ecosystem functioning, the effect of historical legacies relative to other ecosystem properties remains unclear.

Here we estimated how three ecosystem properties – microbial diversity and abundance (hereafter "community structure" as defined by Bier et al. 2015), present-day environmental conditions, and historical legacies – influenced ecosystem functioning in lake sediments. Our approach advanced the search for bivariate B-EF relationships by assessing the importance of diversity in the context of other ecosystem properties. Lake sediments are well suited to test the importance of different ecosystem properties because they: (*i*) share a common microbial species pool from which communities can be differentially assembled according to past events (Niño-García et al. 2016), (*ii*) span large environmental gradients across relatively small distances (i.e. meters), and (*iii*) carry out functions with widespread importance, such as for carbon (C) cycling (Tranvik et al. 2009). Using next-generation sequencing, we first tested for evidence that historical legacies and environmental conditions influenced microbial community composition in lake sediments. We then tested how much *in-situ* organic matter mineralization rates – measured as CO_2 production under ideal conditions – varied with two diversity and two abundance metrics relative to the influence of historical legacies and the present-day environment. CO_2 production is a direct measure of ecosystem function because it is indicative of both food web production and whole-lake C cycling (Tranvik et al. 2009). We predicted that higher levels of diversity, particularly a greater diversity of functional genes, as well as a greater abundance of genes involved in organic matter (OM) decomposition, would increase ecosystem functioning. We also

predicted that microbial community structure would be primarily influenced by variation in environmental conditions rather than colonization history, as expected if dispersal was unlimited ('everything is everywhere, but the environment selects' hypothesis; Baas-Becking 1934). Overall, our results add to limited evidence that the diversity and abundance of both taxa and functional genes explain microbiome functioning in nature, and highlight for among the first time how these effects directly depend on local environmental conditions and indirectly on historical legacies.

2.3 Materials and methods

2.3.1 Study site

We sampled 8 littoral sites each located immediately beneath a discrete catchment drained by a single stream in Daisy Lake, Ontario, Canada (46°270 N, 80°520 W; lake area: 36 ha; maximum depth: 14 m, Fig. A.1). The sites spanned large gradients in the quantity and quality of terrestrial OM inputs from the surrounding vegetation (Tanentzap et al. 2014). These gradients arose from variation in recovery from historical acid and metal contamination, which increased with proximity to a nickel smelter that was closed in 1972 and located 3.5 km northeast of the lake. Following closure of the smelter, lake water chemistry returned to levels characteristic of the broader region: mean ± standard error pH across sites of 6.86 ± 0.02 (Szkokan-Emilson et al. 2011). However, the surrounding vegetation, primarily comprised of paper birch (*Betula papyrifera*) and trembling aspen (*Populus tremuloides*), has been much slower to recover, resulting in a large gradient across sites in terrestrial OM inputs.

Past environmental histories in each site can also result in unique species composition (Martiny et al. 2006). Thus, we considered that the identity of each site would reflect the legacy of historical events with its own unique soils, geomorphology, and dispersal events and these effects would differ from those of the present-day

environment that we could characterize in each site from directly measurable biotic and abiotic variables (described below).

2.3.2 Ecosystem functioning and environmental characterization

We measured ecosystem functioning in 10-15 sediment samples that were collected from each site in waters that ranged from 0.5-1.0 m deep (total n = 97). For each sample, we extruded approximately 35 mL of sediment to a depth of 7.5 cm into 50 mL poly-propylene centrifuge tubes (2.6 cm diameter) using a modified piston corer. Samples were covered with approximately 10 mL lake water from the associated site and transported to the lab on ice.

Ecosystem functioning was measured as total CO_2 production per m^2 after 20 hours. We incubated sediment samples in the dark at 20.5°C and collected headspace gas at the start and end of the incubation by extracting a 2 mL gas sample with an airtight syringe. Gas samples were analyzed on an infrared CO_2 analyzer (Q-S151, Qubit Systems, Kingston, ON, Canada) with a N_2 carrier and converted to mass produced per m^2 using the ideal gas law and a tube surface area of 5.31 cm^2. During sediment sampling, ambient air samples were collected in airtight syringes from each site for subtraction of pre-incubation CO_2 from headspace mass. Sediments were then freeze-dried and stored at -20˚C to stabilize the microbial communities and ensure they were representative of CO_2 measurements (Miller et al. 1999).

We initially characterized sediments with 20 environmental variables. To minimize collinearity, we reduced these to 4 statistically independent predictors that best characterized environmental differences: pH, OM quantity, % terrestrial C, and C:N ratio (see Appendix A.1 and Table A.3 for details). Although this reduction in variables somewhat biased our representation of the local environment, we were primarily interested in estimating the importance of the strongest environmental

gradients relative to other ecosystem properties. First, we measured pH in porewater of each sediment sample at the time of collection with a handheld pH meter (HI9126/HI1230, Hanna Instruments, Woonsocket, RI, USA). Second, 0.5 g of each sediment core was analyzed after incubation for percent OM content as weight loss on ignition (LOI) for 12 hours in a 400°C muffle furnace, confirming the absence of any visible char (Ball 1964). We then used two complementary measurements of OM quality that were averaged at the site-level from a different, unpublished study (Appendix A.1). The first was the percent of terrestrial C in sediment, which characterized the origin of OM. Percent terrestrial C was estimated with a three isotope ($\delta^{13}C$, $\delta^{15}N$, $\delta^{34}S$) mixing model (full details in Appendix A.1). The second measurement was the C:N ratio of sediment, which characterized OM composition. Lower C:N ratios were considered a higher quality to microbial decomposers because they provide more N per mass of sample and are typically associated with more labile material (Taylor et al. 1989).

2.3.3 Microbial communities

We constructed amplicon sequencing libraries for each sediment sample to characterize microbial community composition. Following careful homogenization, DNA was extracted from 0.25 g of each of the samples that had CO_2 measurements using a PowerSoil PowerLyser DNA Isolation Kit (MoBio Laboratories Inc., Carlsbad, CA, USA) according to the manufacturer's instructions. Insufficient fungal sequences were recovered, so we focused on bacteria as representatives of the microbial community. We targeted the V3-V4 region of the 16S rRNA gene using the bacteria-specific 341F-805R primer pair with a two-stage PCR designed for paired-end sequencing. Amplicons were sequenced on an Illumina MiSeq platform (Illumina, San Diego, CA, USA), quality-filtered to remove low-quality bases and putative chimeras, and clustered into

operational taxonomic units (OTUs) at 97% sequence similarity using mothur version 1.39.5 (Schloss et al. 2009). Any read sequenced fewer than six times was removed from subsequent analyses to minimize the influence of spurious reads (Curd et al. 2018). Counts of individual OTUs were then scaled by the total number of reads in each sample to account for sequencing biases using the R package DESeq2 (Love et al. 2014). This measure of "normalized abundance" allows samples with varying read counts to be compared (Knight et al. 2018). Such normalization is widely applied for high-throughput count data (Dillies et al. 2012, McMurdie and Holmes 2014), and all downstream analyses were performed on the DESeq-transformed data to control for these differences in read numbers (Weiss et al. 2017, Knight et al. 2018).

We also constructed shotgun sequencing libraries for 22 of the 97 samples in order to characterize functional genes present in each site (n=2-3 samples per site). Sequencing libraries were prepared with 1 ng of genomic DNA per sample using the Nextera XT DNA Sample Preparation Kit (Illumina) following the manufacturer's instructions and sequenced on an Illumina NextSeq platform. Raw sequences were processed following the EMBL-EBI pipeline version 3.0 (Mitchell et al. 2015) and summarized using Gene Ontology (GO) terms. Sequences were deposited in EBI under project number ERP016063 (full details in Appendix A.1).

Using the microbial sequencing data, we calculated two diversity and two abundance metrics. First, we calculated normalized bacterial abundance by summing the total number of OTUs per sample. The number of OTUs were DESeq-transformed counts rather than relative abundances, so their sum was not equal to 1 and represented differences in normalized abundances between samples (Weiss et al. 2017, Knight et al. 2018). Second, we calculated taxonomic diversity as Shannon's H' for each sample at both the OTU- and the family-level. As both measures were strongly correlated (ϱ = 0.92, $p < 0.0001$), we used the family-level Shannon's H' in our analyses to limit the

number of unclassified taxa whilst retaining as much information as possible about taxonomic diversity (42% of reads were classified to this rank). Shannon's H' is considered a robust estimator of diversity for microbial communities as it accounts for both abundance and evenness (Haegeman et al. 2013), and has been widely used, thereby allowing for comparison with other studies (Ruiz-Gonzalez et al. 2015b, Delgado-Baquerizo et al. 2016). Similarly, we calculated functional diversity on the data obtained from shotgun sequencing using Shannon's H'. Finally, we defined four subsets of functional genes that were involved in different aspects of terrestrial OM decomposition and consequently CO_2 production (after Kirk and Farrell 1987, Sinsabaugh et al. 1994, Golchin et al. 1994, Zhang et al. 2007). We summed the DESeq-transformed abundance of these genes, which were broadly associated with: 1) hydrolase enzymes that break down cellulose, hemicellulose and xylan, 2) oxidases that break down a range of compounds and/or are involved in assimilatory and dissimilatory P and N transformations, 3) intracellular-level carbohydrate metabolism, and 4) aromatic compound catabolism (see Table A.4 for full list of GO categories). We acknowledge that performing GO ontology enrichment tests may have indicated which genes were linked to our function of interest. However, we were primarily interested in testing the hypothesis that the genes most commonly involved in pathways of OM breakdown would be strongly associated with this function. We therefore gave precedent to a hypothesis-based approach rather than one in which we searched for any gene functions potentially involved in OM breakdown.

2.3.4 Is there evidence of legacy and environmental effects on community composition?

We assessed similarity between microbial communities as geographic and environmental distances increased to test if they were associated with historical legacies and present-day conditions. We calculated the Morisita-Horn similarity index for all

pair-wise combinations of normalized microbial abundance, Euclidean distances for geographic distance, and Mahalanobis distances for environmental similarity (i.e. standardized Euclidean distances accommodating for different measurement units and covariance structure among pH, OM quantity, % terrestrial C, C:N ratio). These indexes were calculated using the 97 samples rather than the 22 sample functional gene subset, and significance of the associations between distance matrices was assessed using partial Mantel tests with Pearson's correlation coefficient and 999 permutations constrained within sites. Partial Mantel tests are commonly used to disentangle the effects of present-day environmental conditions on community composition from those of historical legacies, especially when continuous habitat variables and geographic distances are available (Martiny et al. 2006). Here they allowed us to assess the relationship between microbial community and geographic distance while controlling for environmental similarity and vice versa.

We performed a canonical correspondence analysis (CCA) to explore further how bacterial community composition varied in relation to historical legacies and present-day environmental conditions, and in particular, to identify the environmental variables that most explained differences among sites. The CCA was constrained by site, pH, OM quantity, % terrestrial C, and C:N ratio. To test whether community composition varied more with present-day environmental conditions or historical legacies, which we interpreted as being associated with the variation among sites that was unexplained by pH, OM quantity, % terrestrial C, and C:N ratio, we ran a permutational multivariate analysis of variance (PERMANOVA) using the 'adonis2' function in the 'vegan' R package. Significance of marginal effects was assessed with 999 permutations of the community data constrained within sites, with the environmental variables and site identity as predictors. To achieve normality, C:N ratio

was log-transformed while OM quantity (% LOI) and % terrestrial C were logit-transformed.

2.3.5 What is the relative importance of different properties for ecosystem functioning?

We developed a conceptual model to test our hypotheses about the relative importance of community structure, present-day environmental conditions, and historical legacies for ecosystem functioning. This model considered four different pathways by which the different properties could influence ecosystem functioning. Firstly, the model let ecosystem functioning vary with the direct effects of the four environment variables and four measures of community structure (Fig. 2.1). By considering genomic data that could be linked to specific functions, the model also incorporated a trait-based approach that offered more insight into ecosystem functioning than solely based on community-level diversity measures (Krause et al. 2014, Martiny et al. 2015). With the trait-based approach in mind, the normalized abundances of the functional gene sets and measures of OM quality could also interact (dashed arrows in Fig. 2.1), as expected because the efficacy of traits involved in decomposition generally depend on OM quality rather than quantity (Sinsabaugh et al. 2010, Ruiz-Gonzalez et al. 2015a). Thirdly, the model also accounted for the indirect effects that historical legacies, associated with each of the 8 sites, and present-day environmental conditions could have on ecosystem functioning by influencing the four measures of community structure. Finally, we estimated the effects of these four measures on one another to test how the relationship between taxonomy and function indirectly affected ecosystem functioning (small arrows in Fig. 2.1).

We used path analysis to estimate the strength and direction of presumed direct and indirect causal linkages that described our conceptual model (Fig. 2.1). In this analysis, we only used the 22 samples for which we had both taxonomic and functional

information. The model formula for log-transformed CO_2 production as a response in R pseudo-code was:

Ecosystem functioning ~ Environment + Community structure + Functional gene abundance:OM quality + (1|Historical legacies),

where environment included pH, OM quantity, % terrestrial C, and C:N ratio; community structure was each of normalized bacterial abundance, taxonomic diversity, functional diversity, and normalized functional gene abundance; the interaction was between normalized functional gene abundance and each measure of OM quality (% terrestrial C and C:N ratio); and historical legacies were represented by a site-level random effect. We acknowledge that this random effect can also incorporate other present-day environmental variables that systematically varied across sites, but these are unlikely to be more important or uncorrelated with the 20 variables that we actually measured (Table A.1). In total, we fitted the model separately with each of the four normalized functional gene abundance subsets. As gene counts were DESeq-transformed, there was no dependency of one gene on another across samples, and genes could therefore be summed into independent subsets. All of the models also estimated residual (i.e. random) error for each of the focal responses.

We also fitted four separate models to estimate each measure of community structure as a response of the environment, historical legacies, and the other measures of community structure (small arrows in Fig. 2.1). For functional gene abundance, we only modeled the subset(s) of genes identified as significant in the model with CO_2 production as a response. Allowing the measures of community structure to be both dependent variables and independent predictors of ecosystem functioning is consistent with treating them as endogenous variables in a path analysis that teases apart direct and indirect correlations (Grace et al. 2012, Shipley 2016).

Each model described above was fitted with linear mixed models using Bayesian inference by calling the 'blme' function in the R package 'blmer' (Chung et al. 2013). All measured variables were standardized to a common scale with a mean of 0 and an SD of 1, so that we could compare the relative importance of different linkages. To infer effects, we calculated posterior means and 95% confidence intervals (CI) for each parameter by bootstrapping model parameter estimates 800 times with the 'boot' package in R. Effects were considered significant when 95% CI around estimated effect sizes excluded zero. To assess the overall goodness-of-fit of models, marginal R^2 values were calculated.

As all our four models with different functional gene subsets were within 2 small sample Akaike Information Criterion (AICc) units of each other (Burnham & Anderson, 2002; Table A.5), we averaged parameter estimates across the model set. The posterior means of each model were multiplied by their respective AICc weight and summed to determine the average parameter estimates and 95% CI. For functional gene abundance and its interaction with OM quality, no averaging was performed across the model set. We instead reported the effects associated with each of the four unique subsets of functional genes.

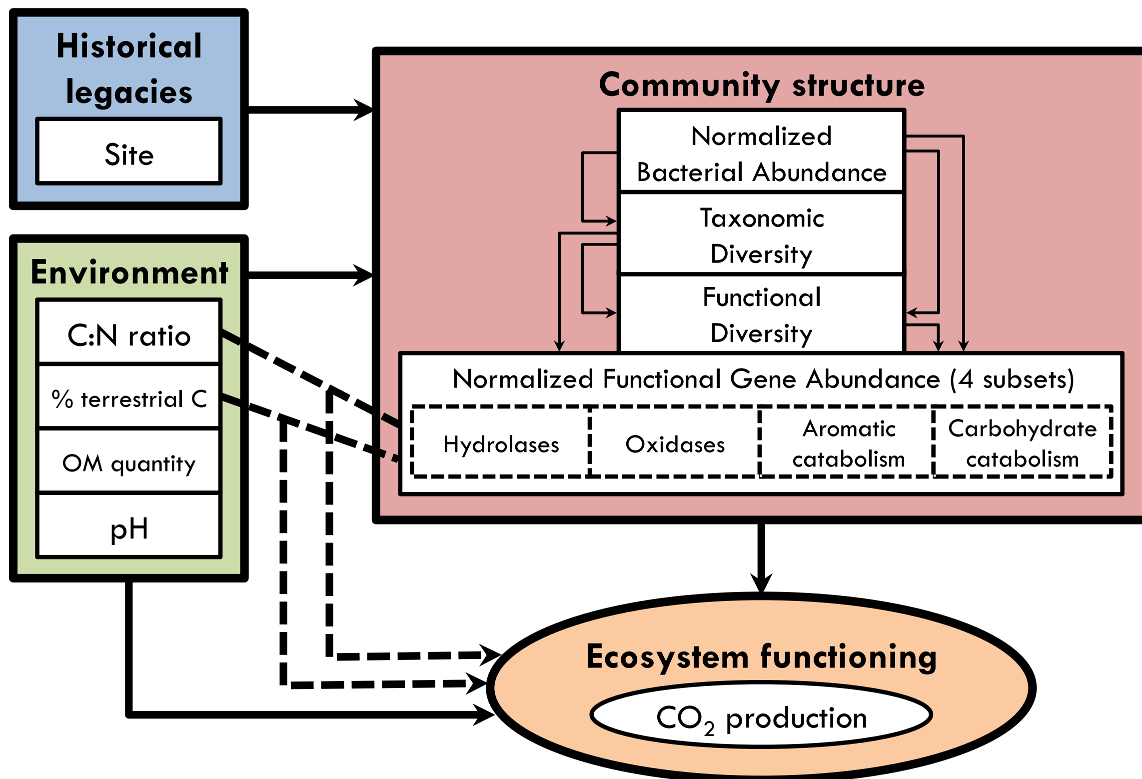

Fig. 2.1 | Trait-based conceptual model of pathways by which different ecosystem properties (community structure, present-day environment, and historical legacies) affect ecosystem functioning. Measured variables associated with different ecosystem properties are in solid white boxes, with the normalized functional gene abundance subsets in dashed white boxes. Solid arrows represent potential pathways between ecosystem properties and dashed arrows represent potential interactions between ecosystem properties.

2.4 Results

2.4.1 Bacterial community composition

Overall, we found considerable bacterial biodiversity. We obtained about 25,000 OTUs that corresponded with about 540 families in each of the 97 and 22 sample datasets (Table A.6). The most common OTUs were in the Koribacteraceae, Hyphomicrobiaceae and Solibacteraceae, each accounting for 0.5% of all normalized abundances per sample. The 22 samples with functional data that we considered in our path analyses subsequently also showed relatively high taxonomic and functional diversity, with H'

values exceeding 5 (Table A.7). The highest number of genes involved in CO_2 production (i.e. normalized functional gene abundance) came from oxidase genes, followed by hydrolases, carbohydrate catabolism, and aromatic compound catabolism (Table A.7).

2.4.2 Biogeographical patterning of the study sites

The 8 sites each showed evidence of having unique bacterial communities that reflected both past events and present-day environmental conditions. We specifically found that communities became less similar as they were increasingly distant in space when controlling for the effects of the environment (partial Mantel test: $r = -0.15$, $p = 0.003$, Fig. 2.2a), suggesting that isolation by distance may maintain differences in biotic assemblages that arose from past events such as differential establishment and persistence of species through time. These differences could not be attributed to dispersal-limitation as most taxa were ‘everywhere’. 369 of the 551 bacterial families occurred at all eight sites, with another 66 present at seven sites (Fig. A.2), evidence that abundances rather than presence varied across space. We also found that community similarity increased with environmental similarity when controlling for geographical distances (partial Mantel tests: $r = 0.43$, $p = 0.005$, Fig. 2.2b), suggesting that different present-day environments also influenced microbial communities within our study.

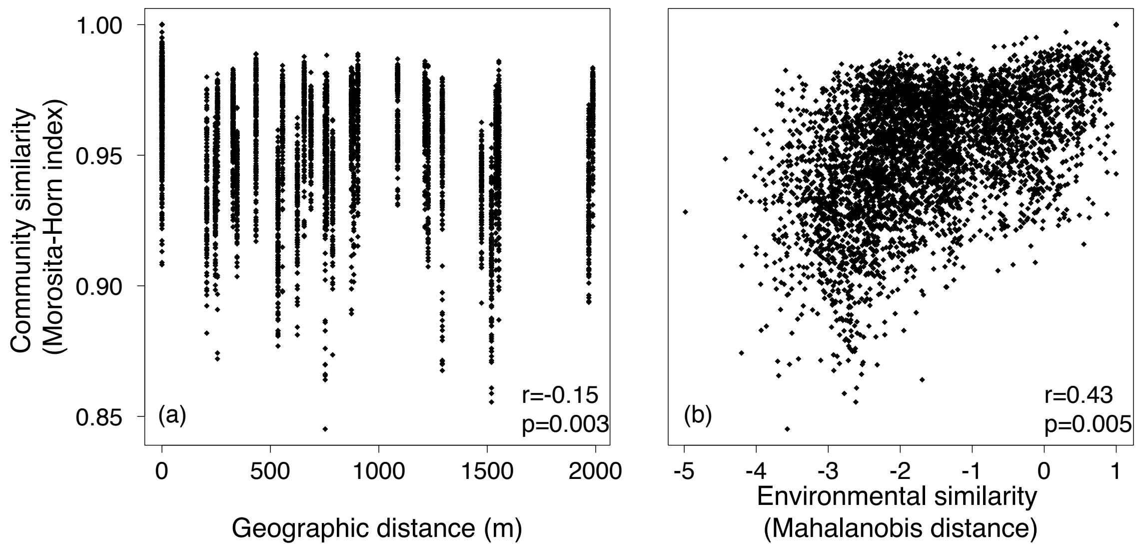

Fig. 2.2 | Bacterial communities differ across geographic and environmental space. Community similarity (Horn-Morisita index) at the family level between all pair-wise combinations of 97 sediment samples (a) decreased with geographic distance and (b) increased with environmental similarity (Mahalanobis distance) according to partial Mantel tests.

We found further evidence of past and present-day influences over bacterial communities when clustering compositional differences among sites (Table A.8a). Distinct communities were observed across sites (F = 3.92, p = 0.01), even after constraining composition by present-day environmental variables (Fig. 2.3). Both measures of OM quality significantly differentiated communities (F = 5.64, p = 0.001 for % terrestrial C and F = 4.61, p = 0.001 for C:N ratio), with no effect of either pH or OM quantity (Table A.8b). We reached a similar conclusion when using a partial redundancy analysis to compare the effects of site identity and environmental variables on community composition (Table A.8c).

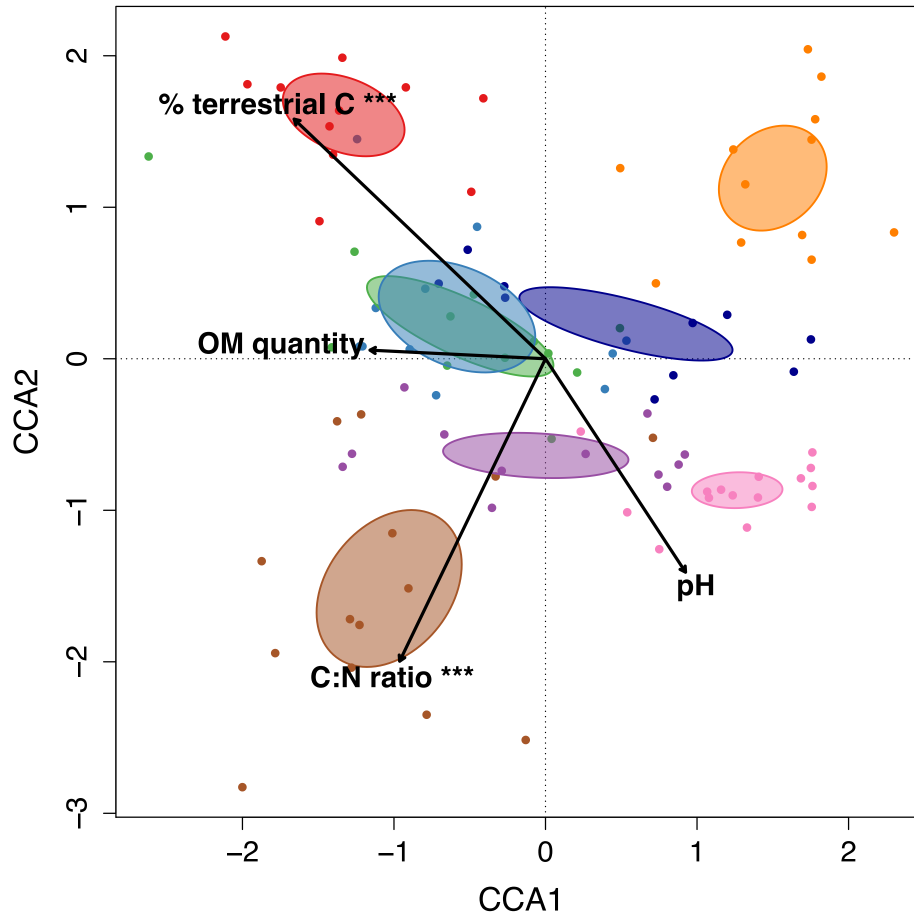

Fig. 2.3 | Bacterial communities differ across sites and environments. The CCA plot shows associations between bacterial community composition at the family level when constrained by site and environmental variables (n = 97). Each color is a distinct site with ellipses representing the standard error around the centroid. Arrows show vector fitting of the constrained environmental variables. * associated with a variable at p <0.05 in the PERMANOVA.

2.4.3 Linking biogeography and community composition to ecosystem functioning

We found that the community structure explained as much variation in ecosystem functioning [median (95% CI): 26% (16-33%)] as the present-day environment [20% (13-29%)], revealing that other properties in addition to those of microbial communities make relatively large contributions to ecosystem functioning (Fig. 2.4). Half of the variation in community structure was attributable to taxonomic diversity (Fig. 2.4). Subsequently, a 1 standard deviation (SD) increase in taxonomic diversity above its

mean H' of 4.4 was sufficient to increase CO_2 production by 77 (95% CI: 36-127) mg m^{-2} from an average of 118 (102-135) to 195 (134-285) mg m^{-2} (Fig. 2.5a). The other half of the variation in ecosystem functioning explained by community structure came from functional genes encoding for oxidases, which were the only gene subset with a statistically significant effect (Tables A.5 and A.9). For example, a 1 SD increase in normalized oxidase abundance above its DESeq-normalized mean of 56.1 increased CO_2 production by 93 (2-317) mg m^{-2} from an average of 139 (99-190) to 232 (101-507) mg m^{-2} (Fig. 2.5b).

An additional 12% (median, 95% CI: 2-18%) of variation in ecosystem functioning was explained by the interaction between community structure and the environment (Fig. 2.4). We specifically found that oxidases further increased CO_2 production when terrestrial C was relatively abundant in sediment, highlighting the dependency of some functional genes on specific environmental conditions for influencing ecosystem functioning. For example, a 1 SD increase in terrestrial C at the mean oxidase abundance increased CO_2 production from the average of 139 mg m^{-2} by 88 (95% CI: 46-154) mg m^{-2} (Fig. 2.5b).

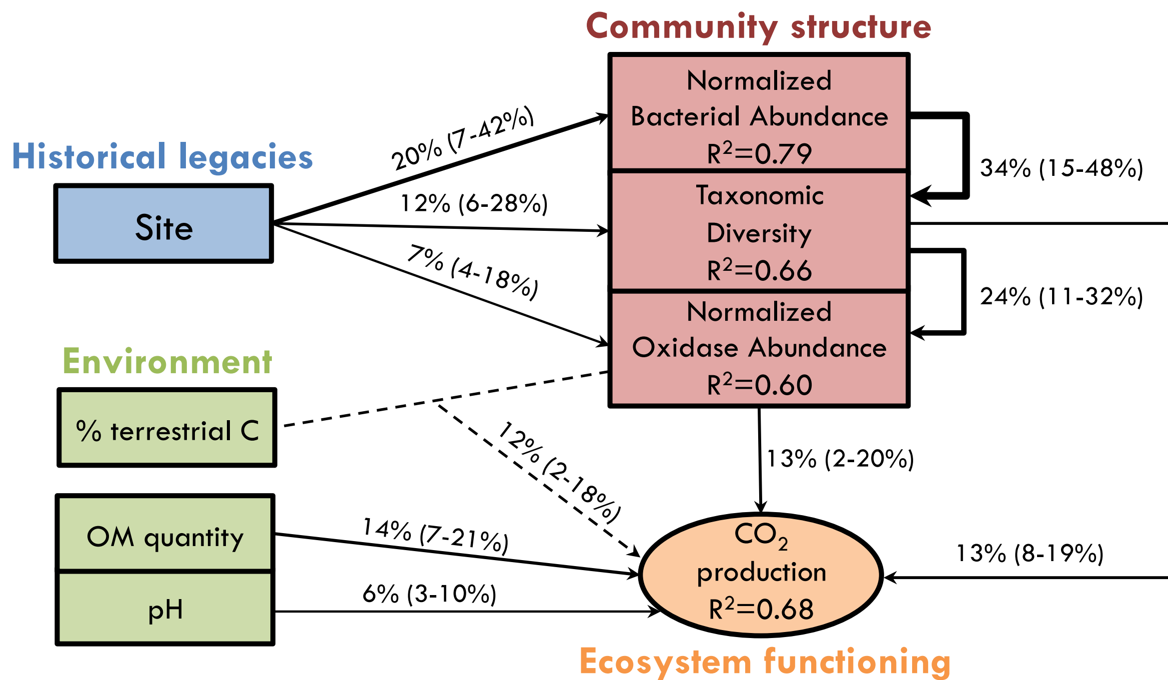

Fig. 2.4 | Individual and interactive effects of the environment and microbial communities largely explain ecosystem functioning. Predictors were averaged across the model set except for normalized functional gene abundance, where oxidases were the best supported subset (Table A.5). Boxes are shown only for variables with a direct or indirect effect on ecosystem function with 95% CI that exclude zero. Numbers accompanying each arrow are median (95% CI) percentage of variance in the associated response explained by a focal effect, with arrow width proportional to these values. Dashed lines represent interactions.

Finally, the averaged model predicting CO_2 production showed that both OM quantity and pH were the strongest environmental correlates of ecosystem functioning in our lake sediments (Fig. 2.4, Table A.9). For example, if OM quantity doubled above its mean value of 13% across our sediment cores, CO_2 production increased on average (95% CI) by 83 (35-167) mg m^{-2} from its mean of 118 (Fig. 2.5c). By contrast, a 1 SD increase in pH above its mean of 5.9 decreased CO_2 production by 23 (9-31) mg m^{-2} (Fig. 2.5d).

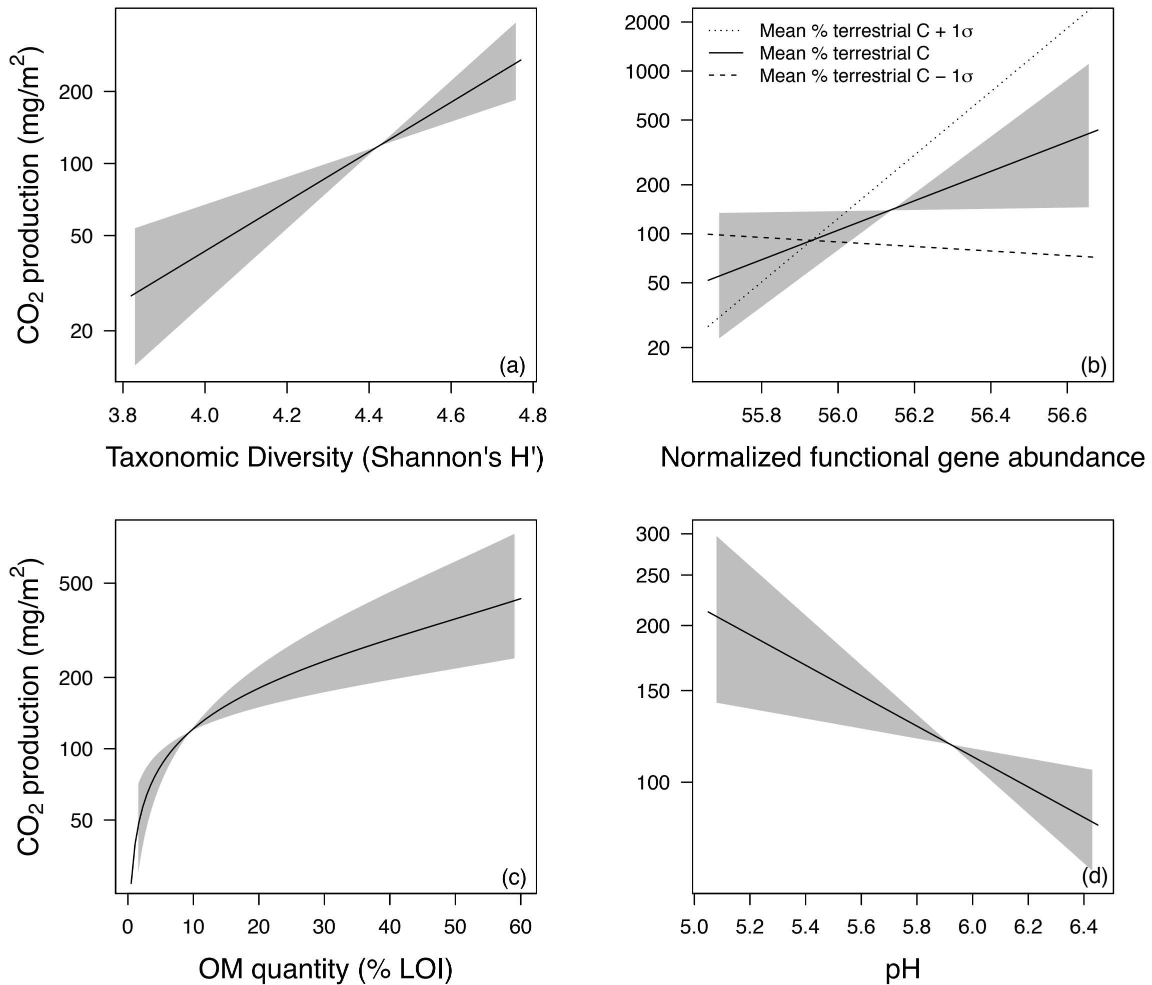

Fig. 2.5 | Environment and community structure influence ecosystem functioning of littoral lake sediments. CO_2 production increases with (a) taxonomic diversity, (b) normalized functional gene abundance (oxidases) as % terrestrial C also increases, and (c) percent OM quantity measured as loss on ignition (% LOI). (d) CO_2 production decreases with pH. Lines are mean model fit at mean values of the other variables. Polygons are 95% confidence intervals. In (b), we show the statistical interaction between normalized functional gene abundance and % terrestrial C by plotting lines at the mean of the latter ± 1 standard deviation (σ).

2.4.4 Indirect effects of historical legacies and present-day environment on ecosystem functioning

The path analysis suggested that the effect of taxonomic diversity in predicting CO_2 production was associated with differences in bacterial communities among sites,

indicative of historical legacies, rather than filtering by the measured environmental variables. A median of 12% (95% CI: 6-28%) of the variation in taxonomic diversity was directly explained by site, with an additional 34% (95% CI: 15-48%), on average, explained by normalized bacterial abundance, which itself was 20% (95% CI: 7-42%) dependent on site (Fig. 2.4). Taxonomic diversity also had an indirect effect on CO_2 production by explaining a median of 24% of the variation in the normalized abundance of oxidase genes (95% CI: 11-32%). This relationship was negative, suggesting that less taxonomically diverse communities were more likely to be dominated by taxa that relied on oxidizing OM as opposed to higher diversity communities where more functions were present (Table A.10). Overall, however, the indirect effects of historical legacies on ecosystem functioning mediated by the community structure were relatively small. Historical legacies explained <20% of the variation in each of the measures of community structure, none of which individually or interactively explained more than 13% of variation in CO_2 production (Fig. 2.4). Thus, even if our site-level random effect included unmeasured present-day environmental variables, these effects were minimal. By contrast, no environmental variables influenced community structure, consequently having no indirect effect on ecosystem functioning (Table A.10). While functional diversity depended on taxonomic diversity, it did not directly influence CO_2 production (Table A.10). We also verified that there were no missing linkages in our model, namely from CO_2 production to community structure, which could feedback onto the latter (Table A.11).

2.5 Discussion

Our study is the first, to our knowledge, that estimates the relative importance of individual and interactive effects of three fundamental properties – community structure, present-day environmental conditions, and historical legacies – on ecosystem

functioning. We found that the present-day environment and community structure explained roughly the same amount of variation in ecosystem function, adding to limited evidence of a positive B-EF relationship in natural microbial communities (Delgado-Baquerizo et al. 2016, Laforest-Lapointe et al. 2017). Importantly, we found that ecosystem processes were predicted better by also considering other characteristics of community structure, like functional gene abundances, and particularly their interactive effects with the environment. These findings, along with the evidence that microbial diversity and abundance varied more with site identity than with the four present-day environmental variables that most differed among sites, highlight the large influence that ecosystem properties other than biodiversity have upon ecosystem functioning.

Our results suggested that, despite their relatively small effects, historical legacies were more important than present-day environmental filtering in explaining the diversity and abundance of microbial communities and thus indirectly influencing ecosystem function. This finding is consistent with others that have shown historical legacies to be important for microbial-mediated ecosystem functions (Hendershot et al. 2017, Martiny et al. 2017). In our study, legacies were likely the result of random differences in colonization history, such as arrival order and timing (Fukami and Morin 2003, Fukami et al. 2010), rather than past geographic events or dispersal limitation, as sites were all located within the same lake and most OTUs occurred at all sites. While measures of OM quality influenced microbial community composition, consistent with others (Schallenberg and Kalff 1993, Ruiz-Gonzalez et al. 2015a), there was no direct effect of these present-day environmental conditions on diversity and abundance, and subsequently ecosystem function. Thus, the biogeographical patterning across sites that we interpreted as being associated with past events adds to growing evidence that, while ‘everything’ may be ‘everywhere’ (Baas-Becking 1934), not everything flourishes

everywhere. One caveat is that there were still many OTUs unidentifiable at the family level.

Contrary to our prediction, we found that environmental conditions directly explained variation in CO_2 production rather than doing so indirectly by changing microbial communities. For example, more acidic samples released more CO_2 due to less inorganic carbon speciation (Morel and Hering 1993). We also found that higher OM quantities increased CO_2 production, potentially because of non-microbial processes, such as extracellular oxidative metabolism and inorganic chemical reactions (Wang et al. 2017). Photo-oxidation or thermal degradation of our samples was unlikely as they were incubated in the dark at controlled temperatures. Alternatively, the increase in CO_2 production could be due to an increase in microbial biomass or in total abundance, but we only estimated normalized microbial abundance in our study. These findings are also consistent with a recent meta-analysis of 58 studies that found microbial biodiversity was not consistently associated with soil environmental variables, but rather more influenced by climate, ecological legacies and evolutionary history (Hendershot et al. 2017). Microbial communities may in part be controlled by processes occurring at finer spatiotemporal scales than routinely measured, e.g. millimeters and minutes (Grundmann and Debouzie 2000).

We found that increasing taxonomic diversity promoted ecosystem functioning more than functional gene diversity, which had no effect on CO_2 production. These results suggest that multiple taxa may perform the same tasks associated with as broad a function as C utilization (Delgado-Baquerizo et al. 2016). Given this convergence of function, increasing numbers of taxa may be sufficient to increase ecosystem functioning irrespective of their specific traits, resulting in high levels of complementarity (Venail and Vives 2013, Laforest-Lapointe et al. 2017). Our results therefore support the need to focus on the identity of traits associated with a response of

interest rather than solely on functional diversity (Krause et al. 2014, Roger et al. 2016, Trivedi et al. 2016). Other metrics, such as phylogenetic diversity, may also enhance predictions of ecosystem functioning by capturing additional axes of trait variation to those directly measured by functional data (Flynn et al. 2011). However, phylogenetic diversity may only be a useful proxy where the associated functions are evolutionarily conserved and not widely dispersed across lineages, as may be the case for a universal function like C utilization.

Oxidase-encoding genes were the only functional gene subset that we found to be associated with ecosystem functioning. Oxidases break down complex and recalcitrant organic polymers that come from terrestrial OM, such as lignin and humic acids (Sinsabaugh 2010), and which would have varied considerably across sites given the surrounding forest gradient (Tanentzap et al. 2014). The increasing association between oxidases and CO_2 production as terrestrial C inputs to sediments increased was also unsurprising as aquatic microbial communities are adapted to utilize complex organic polymers derived from litterfall (Judd et al. 2007, Emilson et al. 2017). We also found more oxidase genes in less taxonomically diverse assemblages, which may have arisen if a few taxa containing a higher proportion of oxidase genes became disproportionally active with increasing inputs of terrestrial C (Muscarella et al. 2016). The other subsets of genes may have not explained much variation in CO_2 production because they affected simpler molecules that were less associated with our specific terrestrial C gradient (Rocca et al. 2015). Additionally, shotgun metagenomics can only ascertain the presence of genes, not their expression, so some gene sets may have had nonsignificant effects because they were inactive.

Our results also suggest that future increases of OM inputs may promote benthic respiration and reduce the large C sink capacity of many northern lakes (Gudasz et al. 2017). Northern lakes are burying increasingly more terrestrial OM into their sediments,

primarily due to human activities (Heathcote et al. 2015). Across our sites, a difference in sediment OM of 2 vs 55%, associated with a doubling of surrounding forest cover from 36% to 64% (Szkokan-Emilson et al. 2011), was sufficient to increase CO_2 production by an average (95% CI) of 5-times (3-7 times). Thus, our results also show how models that integrate biodiversity and trait-based approaches can better predict the outcomes of future changes to lake C cycles. More broadly, predictions of how ecosystem functioning varies with biodiversity will be improved if placed in the context of other ecosystem properties, such as past and present-day environments. Future studies should consider generalizing the importance of these other properties relative to the taxonomic and functional aspects of biodiversity in different spatial and temporal contexts. Another next step from our study would be to disentangle past and present-day influences more directly, such as by manipulating colonization dynamics in different environmental contexts and measuring how ecosystem function responds (e.g. Reed and Martiny 2013).

Chapter 3

Biotic and abiotic mechanisms independently stabilize different measures of ecosystem functioning

3.1 Abstract

Species diversity can stabilize ecosystem functioning, such as by directly increasing the mean or reducing the variance of a function or indirectly from compensatory fluctuations in species abundances. Yet few studies outside of primary producers have simultaneously tested for how the strength and direction of diversity-stability (D-S) relationships and their underlying mechanisms are influenced by environmental changes that modify niche and resource availability. Here we tested support for the D-S relationship and its underlying mechanisms in lake sediment microbial communities, which carry out functions of widespread importance and span large environmental gradients. We simulated future environmental changes by creating nearshore sediments with different terrestrial organic matter (t-OM) quantity and quality, and measured ecosystem function over one year as CO_2 production and microbial abundance, indicative of microbial activity and growth. While diversity stabilized CO_2 production by reducing its temporal variation, this effect did not vary along the environmental gradients. However, higher t-OM inputs stabilized microbial abundances regardless of diversity levels by promoting community-level species asynchrony that reduced

temporal variation in ecosystem functioning. These results contrast other systems where increases in the temporal mean of functions are more important. More broadly, our study reveals that whether environmental changes influence the stabilizing effect of diversity on ecosystem functioning will depend on the function in question.

3.2 Introduction

Diversity stabilizes ecosystem functioning in many ways (Tilman 1999), but little is known about when different mechanisms operate. Stability is routinely measured as the ratio between the mean μ and variance σ in ecosystem functioning over time (Tilman 1999). Much of our understanding of positive diversity-stability (D-S) relationships comes from experiments measuring the productivity of primary producers (Shurin et al. 2007, Boyer et al. 2009, Downing et al. 2014, Hautier et al. 2014, Jucker et al. 2014, Ramus and Long 2016, del Río et al. 2017). Yet few have tested whether positive D-S relationships hold outside of the laboratory at other trophic levels, such as microbial decomposers, and even fewer have tested its potential causes in these communities (Downing et al. 2014, Wagg et al. 2018). As global change is altering species diversity (Cardinale et al. 2012), understanding when the stabilizing effect of diversity on ecosystem functioning will vary can help improve the delivery and predictability of ecosystem processes (McGrady-Steed et al. 1997, McCann 2000).

Environmental change can ultimately influence the strength and direction of D-S relationships through at least three different proximal mechanisms. First, as new niches become available, species can partition resources and minimize inter-specific competition, promoting diversity and causing communities to have higher mean μ function (Hector et al. 2010). This mechanism, called overyielding, can consequently stabilize functioning if environmental changes increase niche availability (Micheli et al. 1999, Downing et al. 2008). Second, ecosystem functioning can be stabilized as

diversity increases in the absence of new niches if its variability σ is reduced. We term this mechanism undervarying. Undervarying can arise from the partitioning of existing niches, such as by shifts towards facilitative rather than competitive interactions, which subsequently enhance temporal coexistence (del Río et al. 2014, Hunting et al. 2015). Again, any environmental change that influences species coexistence may change σ and thus D-S relationships via an undervarying mechanism. Finally, more diverse communities can maintain ecosystem function because there are more species that can increase in abundance and compensate for declines in the abundance of other species that contribute to functioning (Houlahan et al. 2007, Hector et al. 2010). This mechanism of species fluctuating asynchronously will again be influenced by changes in resource availability and the different responses of species to environmental stress (Ives et al. 1999, Thébault and Loreau 2005, Loreau and de Mazancourt 2008). Yet few studies have simultaneously tested for these stabilizing mechanisms and their environmental dependency outside of primary producers (Grman et al. 2010, Downing et al. 2014).

Here we tested whether support for the D-S relationship and its underlying mechanisms varied along two environmental gradients. We carried out our study in lake sediment microbial communities because they carry out functions of widespread importance, such as carbon (C) cycling, and span large environmental gradients across small distances, i.e. meters (Orland et al. 2018), which can influence their functioning. Additionally, inputs of terrestrial organic matter (t-OM), especially aromatic and recalcitrant compounds, are expected to increase in northern waters (Creed et al. 2018), and are likely to change the support for D-S by modifying the local environment. We therefore simulated these future environmental changes by creating nearshore sediments with different t-OM quantity and quality, and measured ecosystem function over one year as CO_2 production and microbial abundance, indicative of microbial activity and

growth (Nielsen et al. 2011). We then tested whether microbial diversity and t-OM inputs – individually and interactively – increased the stability of either function through: 1) *overyielding*: increasing the function μ; 2) *undervarying*: reducing the function σ; and 3) *asynchrony*: temporal complementarity in species abundances that increases μ and/or reduces σ.

We predicted that diversity would ultimately stabilize ecosystem function but through different proximal mechanisms. First, as CO_2 production is a ubiquitous function (Carlson et al. 2007, Wertz et al. 2007), we expected greater diversity would consistently be associated with greater mean CO_2 production and thus ecosystem stability, i.e. via overyielding. Second, greater diversity should increase the chance that some species can incorporate t-OM into their biomass and lead to differences in abundances that enhance species asynchrony. We therefore expected that, as t-OM increased, diversity would be more strongly associated with asynchrony, which would stabilize the abundance of the overall community because of compensatory dynamics among species. More t-OM should also increase community stability independently of diversity because t-OM is primarily allocated to biomass (Guillemette et al. 2016), especially in dark lakes as ours (Fitch et al. 2018), and so should be directly associated with changes in abundances.

3.3 Materials and methods

3.3.1 Study design

We submerged experimental mesocosms on the bottom of Lake Laurentian, Canada (46°27’30” N, 80°56’0” W) beneath 0.30-0.75 m of water during July 2015. Lake Laurentian is a small lake (1.57 km^2 area) surrounded by early-successional boreal forest and with minimal human disturbance. The lake is generally mixotrophic as per Williamson et al. (1999) because it has relatively dark waters (colored dissolved organic

carbon concentrations measured as absorbance at 320 nm of 26 m^{-1}) and high average (± standard error) total phosphorus concentrations of 35.2 (± 2.5) $\mu g\ L^{-1}$ based on summer mid-lake surface grabs.

Mesocosms were filled with ca. 15 L of sediment consisting of different types of t-OM and housed in HDPE containers (surface area: 0.19 m^2, depth: 0.13 m) after Tanentzap et al. (2017). Briefly, we added 5, 25, and 50% t-OM on a dry-weight basis to locally sourced inorganic material with particle sizes and vertical structuring of all material mimicking natural lake sediments (Tanentzap et al. 2017). For each t-OM quantity, material was comprised of 33%, 50%, or 66% deciduous litterfall (primarily *Acer rubrum*, *Betula papyrifera*, *Populus tremuloides*, *Quercus* spp.) by dry-weight collected from nearby forests. Coniferous litterfall (*Pinus* spp.) comprised the remaining material. Each treatment was then replicated three times, resulting in 3 t-OM quantities × 3 t-OM qualities × 3 replicates. We also included a control treatment filled with inorganic material only (total n = 30). Mesocosms were arranged in a block design between two sampling bays, submerged in rows, and covered with a 1 mm × 1 mm nylon mesh screen to standardize the percentage of sunlight reaching the sediment surface. After one month, we made an 8 cm slit in the center of each screen to collect sediment. To collect porewater, we secured a 3 mL polypropylene syringe horizontally immediately beneath the sediment surface. The wall of the syringe that faced the sediment was removed and covered in ca. 250 μm nylon mesh. Importantly, sediment porewater samples taken from our mesocosms reflect the biogeochemistry of the surrounding natural lake sediment (Tanentzap et al. 2017), allowing us to extrapolate our findings to field conditions.

3.3.2 Stability and diversity measures

We calculated the stability of two ecosystem functions – CO_2 production and microbial community abundance – by dividing the temporal mean of each function μ by its standard deviation σ (after Tilman 1999). Data collection began one month after mesocosms were in-
stalled to ensure that the treatments had settled.

CO_2 stability was calculated from approximately fortnightly measurements of porewater during ice-free conditions between August 2015 to September 2015 and May 2016 to August 2016 (n = 18 dates per mesocosm). On each sampling occasion, we extracted 43 mL of porewater into a 60 mL syringe that was pre-acidified with 2 mL of 0.5M HCl. 15 mL of atmospheric air was pulled into the syringe, which was shaken for 2 minutes and left to equilibrate for 30 seconds. 10 mL of the headspace was then analyzed for CO_2 on a SRI 8610C-0040 gas chromatograph (Torrance, CA, USA) within 24 hours of collection. Porewater concentrations were calculated by subtracting ambient air additions and applying the Bunsen solubility coefficient and ideal gas law (Aberg and Wallin 2014), accounting for porewater pH and temperature simultaneously measured in the field with a handheld meter (HI 9126, Hanna Instruments, Woonsocket, RI, USA). To achieve normality, CO_2 measurements were log-transformed prior to calculating stability.

Microbial community stability was estimated from the total abundance of bacterial operational taxonomic units (OTUs) in each mesocosm between two sampling occasions in September 2015 and August 2016. OTUs were identified from surface sediment (ca. top 5-cm) grabs that were sequenced on an Illumina NextSeq (300 cycles, paired-end) after all environmental DNA was extracted (Appendix B.1). Raw sequences were processed at a depth of approximately 3.3 million reads per sample following a modified version of the European Molecular Biology Laboratory-EBI pipeline version

3.0 (Mitchell et al. 2016) and were taxonomically-annotated with representative 16S sequences using QIIME v.1.9.1 (Appendix B.1). The sequences were deposited in EBI under the project accession number ERP019980. From all samples, we removed singletons and 14 OTUs (out of the remaining 7,697) that were present in a negative water-only control with a relative abundance of >1%. Differences in the number of reads per sample due to sequencing biases was accounted for using a variance stabilizing transformation with the R package DESeq2 (Love et al. 2014). As the number of OTUs were DESeq-transformed counts rather than relative abundances, they did not sum to 1 and their sum approximated the total abundance of each sample (Weiss et al. 2017).

We also calculated two other measures of microbial community structure. First, we calculated microbial diversity by averaging the total number of unique OTUs (i.e. richness) between the two sampling periods in each mesocosm. Second, we calculated community-wide asynchrony *A*, which estimates how individual species, or OTUs, differentially fluctuate in their abundances over time. Although individual OTUs may be transient, communities may be stabilized if the increase in a taxon is compensated by the decrease in another. Here, we calculated *A* for each mesocosm by comparing the temporal variance in the summed OTU abundances in each sample MA_{σ}^{2} with the summed variances of each OTU i in each mesocosm with n OTUs (Loreau and de Mazancourt 2008, Hautier et al. 2014):

$$A = 1 - \frac{MA_{\sigma}^{2}}{(\Sigma_{i=1}^{n} MA_{\sigma i})^{2}}$$

A varies between 0 (no asynchrony) and 1 (perfect asynchrony).

3.3.3 Does diversity stabilize ecosystem function as the environment changes?

We tested whether microbial diversity stabilized each of our two measures of ecosystem function under changing environmental conditions using linear models. The two measures were weakly inter-correlated, as expected if they captured distinct functions (Spearman's correlation comparing temporal means and stability between each function: ϱ = 0.13 and 0.08, respectively). Our models subsequently included the following predictors: diversity, two measures defining our environmental gradient in terms of t-OM quantity and t-OM quality, and asynchrony, and we added interaction terms between diversity and each environmental variable. Diversity and the environment can ultimately influence stability through three proximate mechanisms: overyielding, undervarying, and asynchrony. However, we included asynchrony in the models as it influences stability directly rather than its components μ and σ like the former two mechanisms. Sampling bay was also included to account for the blocking design of our experiment. Diversity and stability measures were log-transformed, and all measured continuous variables were standardized to a common scale with a mean of 0 and a standard deviation of 1 to compare their effects. If we found a statistically significant effect of asynchrony, we fitted an additional linear model to test whether this arose because asynchrony itself was ultimately influenced by individual or interactive effects of diversity and the environment.

3.3.4 Which proximal mechanisms stabilize ecosystem function?

We tested other proximal mechanisms for the estimated D-S relationships in addition to asynchrony. If either diversity or the environment stabilized ecosystem functioning in our previous model, we tested whether these effects arose because of increasing μ (i.e. overyielding) and/or decreasing σ (i.e. undervarying) by fitting separate linear models to each variable with the same predictors as for stability (excluding asynchrony). We also

tested whether increased stability in more asynchronous communities resulted from a higher mean and lower variation in total microbial abundance by modeling each of these variables as a response of asynchrony and sampling bay.

3.4 Results

3.4.1 Diversity stabilizes ecosystem function regardless of the environment

We found a stabilizing effect of diversity on CO_2 production ($t_{22} = 2.18$, $P = 0.041$; Fig. 3.1). This effect did not vary with the environment and neither t-OM quantity nor quality were directly associated with stability on their own (Table B.1).

While diversity stabilized CO_2 production, we found no direct effect of diversity or the environment on community stability (Table B.1). Sampling bay had no effect on either measure of ecosystem stability (Table B.1).

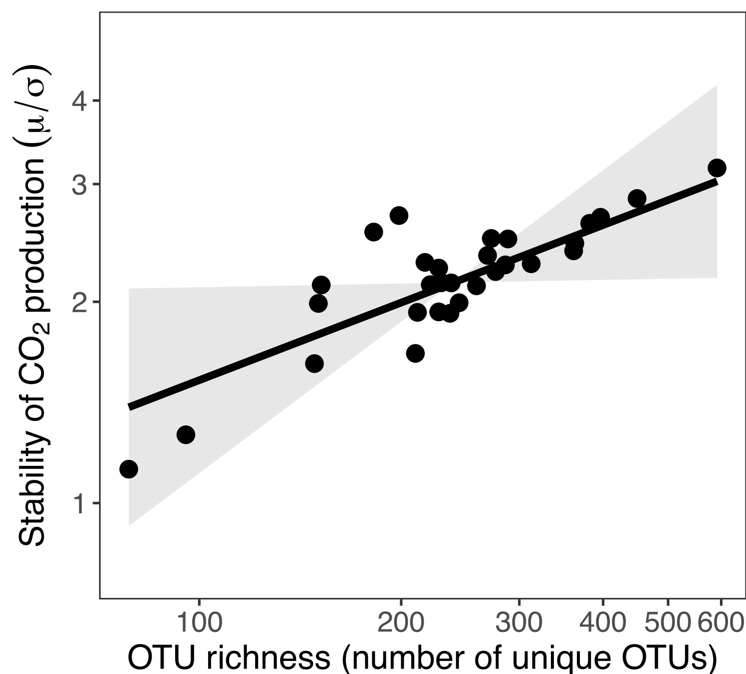

Fig. 3.1 | Mean OTU richness stabilizes CO_2 production. Line is mean model fit at mean values of the other predictor variables. Polygon is 95% confidence interval.

3.4.2 Undervarying and asynchrony stabilize ecosystem function

The stabilizing effect of diversity on CO_2 production was the result of reduced temporal variation consistent with an undervarying mechanism (t_{23} = -2.60, P = 0.016). By plotting the slope estimates of diversity separately regressed against the μ and σ in CO_2 production, we found that the effect of OTU richness on the latter metric was more negative than the reduction in mean CO_2 production with increased diversity (i.e. effect diverged from the 1:1 line, Fig. 3.2). We found no evidence for overyielding in our study lake (Table B.1).

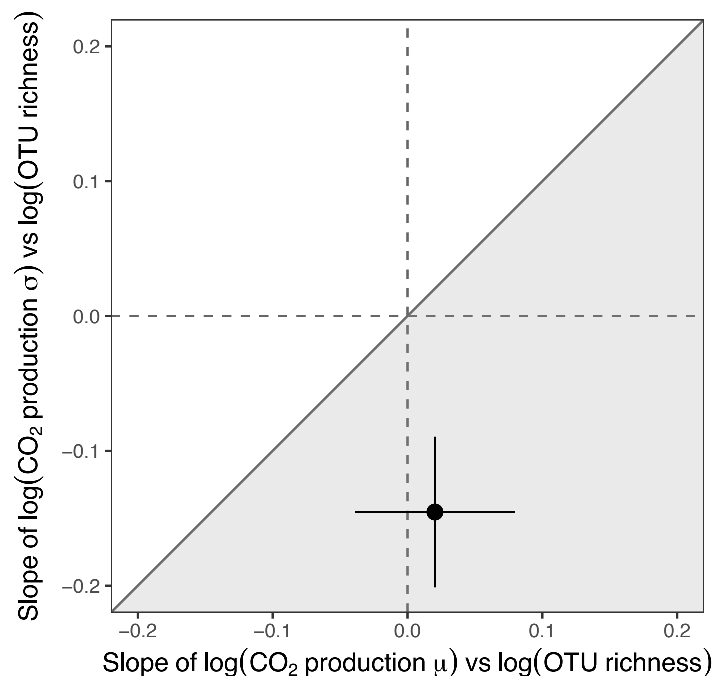

Fig. 3.2 | Undervarying can explain the stabilizing effect of diversity on CO_2 production. Point is mean ± SE for slope of the temporal mean CO_2 production μ vs diversity compared to the slope of the temporal SD in CO_2 production σ vs diversity. Grey area of the plot (below the 45° line) indicates a stabilizing effect arising from overyielding (points to the right of the vertical dashed line) and/or undervarying (points below the horizontal dashed line).

While there was no direct effect of diversity or the environment on community stability, we found a positive effect of community-wide asynchrony (t_{22} = 3.99, P <

0.001; Fig. 3.3a), consistent with our prediction that, by fluctuating differentially through time, species would compensate for each other and maintain abundances. The effect of asynchrony arose by reducing the σ of total microbial abundance through time in each mesocosm (t_{27} = -4.39, P < 0.001; Fig. 3.3b). For example, a 1 SD increase in asynchrony above its mean of 0.68 increased stability from an average of 1.80 to 3.38 (95% confidence interval for increase: 2.44-4.68). There was no stabilizing effect of asynchrony on the μ of total microbial abundance (Table B.1).

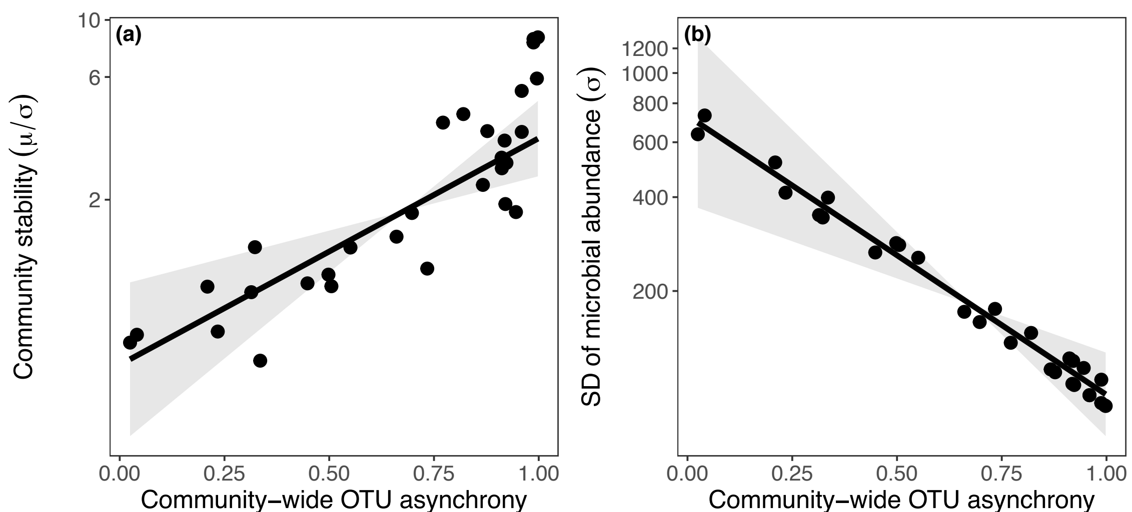

Fig. 3.3 | Asynchrony stabilizes microbial abundance by undervarying. Asynchrony (a) increases community stability by (b) decreasing the standard deviation (SD) of the total microbial abundance in each mesocosm across time. Line is mean model fit at mean values of the other variables. Polygon is 95% confidence interval.

Additionally, we found that asynchrony itself increased with t-OM quantity (t_{23} = 3.06, P = 0.006), but not with diversity (Table B.2), suggesting that the environment indirectly stabilized the community. Whether the t-OM was comprised of deciduous or coniferous litterfall did not influence these results (Table B.2).

Overall, our results support the prediction that microbial diversity ultimately stabilizes CO_2 production regardless of important environmental conditions through an

undervarying mechanism (Fig. 3.4). Yet, we also found that increasing t-OM inputs did stabilize microbial abundances by promoting asynchrony and without necessarily influencing diversity (Fig. 3.4).

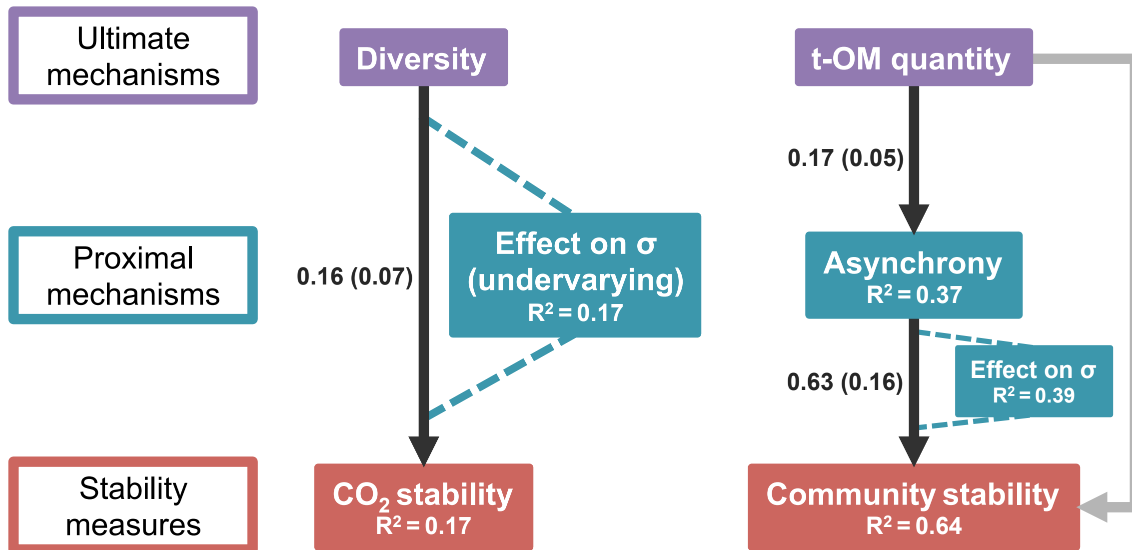

Fig. 3.4 | Diversity directly stabilizes CO_2 production while t-OM quantity indirectly stabilizes microbial community abundance. Ultimate mechanisms, proximal mechanisms, and stability measures are shown in purple, blue, and red, respectively. Significant ($p<0.05$) effects between variables are shown with black lines, potential pathways (non-significant) shown by grey lines. Dashed lines describe proximal mechanisms, i.e. whether functions are stabilized by increased μ or decreased σ (only significant results are shown). Numbers adjacent to each arrow are standardized parameter effects (with standard errors). R^2 shown for all modeled responses.

3.5 Discussion

Our results add to growing evidence for positive D-S relationships in microbiomes (McGrady-Steed et al. 1997, Girvan et al. 2005, Ptacnik et al. 2008, Eisenhauer et al. 2012, Downing et al. 2014, Tardy et al. 2014, Wagg et al. 2018) and also shed light on their underlying mechanisms. We found that both diversity and environmental conditions can stabilize ecosystem functioning and do so via different mechanisms

depending on the ecosystem function. As we anticipated, the D-S relationship did not vary along environmental gradients for a ubiquitous function like CO_2 production, suggesting that more diverse communities can maintain ecosystem functioning across a range of sediment conditions. However, higher t-OM inputs stabilized microbial abundances regardless of diversity levels. Depending on the function of interest, environmental changes may therefore override any stabilizing effect of diversity. The stabilizing effect of the environment was still related to community composition though, as it occurred via an increase in species asynchrony.

Diversity stabilized CO_2 production by reducing its temporal variation rather than increasing its mean, consistent with an undervarying mechanism. Undervarying is more likely to arise from facilitative interactions and greater niche partitioning rather than declines in competition (Mulder et al. 2001, Isbell et al. 2009, del Río et al. 2014), which tend to be associated with an overyielding mechanism (Hector et al. 2010). For example, one species' decomposition products may provide a source of nutrients for another species (Tardy et al. 2014), thus increasing the community's overall metabolic activity. These positive interactions do not imply, however, that the community should grow in size. Functional changes do not always map onto abundances and vice versa (Orland et al. 2018). Moreover, the lack of an effect of t-OM inputs on the D-S relationship was consistent with our prediction that CO_2 production is a widely performed microbial function that will operate across a broad range of environments (Carlson et al. 2007, Wertz et al. 2007). Although we had hypothesized there would be a similarly positive effect of diversity on microbial community stability, measures that we did not consider here could be more important than species diversity, such as differential species establishment, colonization history, and phylogenetic diversity (Sankaran and McNaughton 1999, Cadotte et al. 2012).

Sediment conditions – specifically high levels of t-OM – stabilized microbial abundances by increasing community-level species asynchrony regardless of diversity. At lower t-OM quantities, taxa may be resource-limited and all more likely to decrease consistently in abundance, thus decreasing community-wide asynchrony. By contrast, there may be a greater fluctuation of taxa at higher t-OM concentrations (Tilman 1999), because some taxa will be able to use t-OM more effectively than others (Findlay 2003, Berggren et al. 2010b). Greater t-OM may also promote competition and increase variation in the abundance of individual species and their asynchrony (Tilman et al. 1998, Loreau and de Mazancourt 2008). Unsurprisingly, considering asynchrony has been shown to saturate quickly with increasing species richness (Jucker et al. 2014), we found no association between diversity and asynchrony in our relatively species-rich communities.

Our results are consistent with theory that asynchrony is a strong predictor of stability (Loreau and de Mazancourt 2013) and that its effects are largely driven by environmental changes (Ives et al. 1999, Thébault and Loreau 2005, Downing et al. 2008). While there is evidence for a relationship between asynchrony and stability in plant and phytoplankton communities, we provide among the first evidence that asynchrony can stabilize bacterial communities in nature (Wagg et al. 2018). Compensatory dynamics are expected to be more important in communities with rapid species turnover like microbial communities, as microbes have short generation times and can thus readily fluctuate in numbers. Asynchrony will also be enhanced in highly diverse communities, where the addition of new species will provide a form of ecological insurance that functioning will be maintained (Yachi and Loreau 1999, Tardy et al. 2014). This idea is consistent with our observed negative association between asynchrony and temporal variation in community abundance. However, as functionally similar species are expected to respond similarly to environmental changes,

communities that contain more functionally similar species are expected to be less asynchronous (Hector et al. 2010, Roscher et al. 2011). This may explain why we did not observe an effect of asynchrony on the stability of CO_2 production, a function shared widely among microbes.

We found no evidence for overyielding as a mechanism stabilizing CO_2 stability or community abundance. More diverse communities did not have greater mean microbial abundance, indicating that inter-specific competition may reduce the abundances of individual species despite more species being present. Different processes were therefore at play depending on the function in question: competition for microbial abundances, and facilitation for CO_2 stability, as described above. More species-rich communities also did not produce more CO_2, suggesting that taxonomic diversity is not necessarily associated with greater ecosystem functioning. These results are unsurprising as high functional redundancy and incomplete taxonomic resolution can obscure positive biodiversity-ecosystem functioning relationships in microbial communities (Nielsen et al. 2011, Martiny et al. 2015). These relationships may also depend on specific environmental conditions or historical events that structure these communities (Orland et al. 2018). Finally, overyielding may be more important in single trophic level systems like grasslands than in complex microbial communities (Downing et al. 2014). For example, an increased diversity of predator species can increase mean community respiration but this effect can be offset by a corresponding reduction in prey species.

North temperate lakes are undergoing rapid environmental changes, characterised by increased loadings of terrestrial organic material (Creed et al. 2018), and our results suggest that these changes will increase microbial community stability by promoting asynchronous fluctuations in species abundances. These results are supported by previous work showing that terrestrial organic carbon stabilizes bacterial

communities in experimental ponds (Muscarella et al. 2016). Our study now advances this work by showing the mechanisms that underlie these responses. While our findings also suggest that future environmental changes will not alter the stability of CO_2 production, stabilized microbial abundances may maintain other functions besides CO_2 production (Wagg et al. 2018), like nutrient cycling (Finlay et al. 1997). Changes to the environment unrelated to t-OM inputs and unmeasured in our study may also affect CO_2 stability if they reduce species diversity, such as nutrient loading and contaminant deposition (Paerl et al. 2002, Zeglin 2015). Overall, predicting dynamics of microbial communities is difficult both because of the complexity of biotic interactions and the adaptive capacity of species that allows them to evolve with their environment (Magurran et al. 2010, Thomas et al. 2018). Our study nonetheless identifies why the important functions undertaken by freshwater microbes can be maintained through time, thereby helping to predict future responses to a changing environment.

Chapter 4

Think global, act local: small-scale environmental filters primarily influence microbial community development

4.1 Abstract

The early stages of community development are important for determining the species and traits that establish within communities. Both small- and large-scale abiotic conditions can filter species assemblages at different stages of community assembly, but little is known about the scale-dependency of environmental filtering in microbes. Here we tested how different environmental scales influenced the rate at which microbial community composition changed over time in lake sediments, and whether these changes occurred synchronously across different environments given the same initial communities. We manipulated the small-scale environment by creating sediments with different terrestrial organic matter (t-OM) quantity and quality and placing these in three lakes differing in trophic status and representative of the large-scale environment. We then characterized microbial communities over a two-month period and found that communities became more dissimilar with time despite being derived from the same initial leaf material. Our results revealed that small- as opposed to large-scale environmental conditions were most important for filtering microbial community composition in contrast to findings from macro-organisms. Sediment t-OM quantity

was the primary small-scale filter that explained dissimilarities between mesocosms and the only predictor of connectivity among taxa. Additionally, microbial kingdoms varied by up to 10-times in how quickly they changed, providing among the first evidence that they respond differently over time to their surrounding environment and/or biotic interactions. Our analysis highlights that future changes to both sediments and lake waters can modify how lake sediment microbial communities develop with far-reaching consequences for important ecosystem functions like carbon cycling.

4.2 Introduction

Understanding how communities develop with time is important because it influences ecosystem functioning. Microbial communities offer a good system to study compositional changes in real time as they assemble and turnover across relatively short timescales, i.e. days (Hewson et al. 2006, Redford and Fierer 2009). Most of what is known about how microbial communities assemble in nature comes from soils, glaciers, and biofilms (e.g. Jackson et al. 2001, Martiny et al. 2003, Nemergut et al. 2007, Schütte et al. 2010, Fierer et al. 2010, Dini-Andreote et al. 2015, Smith et al. 2015), with much less known in lake sediments that perform many important ecosystem functions (Tranvik et al. 2009). Lake sediments differ from systems like soils because they are much more connected to the surrounding landscape through flow pathways, so should respond differently to small- and large-scale abiotic conditions that determine which species establish and persist in a site, i.e. environmental filtering (Langenheder and Székely 2011, Kraft et al. 2015). Changes in these communities may also occur at different rates across these environmental scales (Kent et al. 2007).

Little is known about how the rate at which communities develop with time changes with small- and large-scale filters, especially for lake sediment microbes that perform important ecosystem functions. In plant communities, these filters appear

sequentially with time. Composition is initially selected by the adaptation of species to large-scale environmental conditions such as climate (Wiens and Donoghue 2004, de Bello et al. 2013). Small-scale filters such as soil characteristics then act on abundances and co-occurrence patterns (Woodward and Diament 1991, Diaz et al. 1998, Chesson 2000). In microbes, community assembly and biogeography have been extensively studied (Martiny et al. 2006, Tedersoo et al. 2014, Delgado-Baquerizo et al. 2018a), but not how rates of community change might vary with environmental filters that arise at different spatial scales. Environmental filters have simply been identified at both a large-scale, such as connectivity among catchments and lake conditions (Nelson et al. 2009, Nino-Garcia et al. 2016), and smaller scale, such as resource quantity and quality (Ruiz-Gonzalez et al. 2015a, Tanentzap et al. 2014) and pH gradients (Fierer and Jackson 2006). In parallel, microbial community composition, primarily in bacterial biofilms, has been found to change predictably over time (Jackson et al. 2001, Martiny et al. 2003, Lyautey et al. 2005, Redford and Fierer 2009). However, there has been no attempt, to our knowledge, to test how different environmental filters influence temporal changes in lake sediment microbial communities.

How species co-occur will also influence how lake sediment microbial communities change with time (Gotelli and McCabe 2002). Co-occurrence patterns can arise because of shared environmental preferences or biotic interactions (Leibold and McPeek 2006, Fuhrman and Steele 2008). These co-occurrence relationships are especially important as they may alter the stability of a community over long timescales (HilleRisLambers et al. 2012). A species with reduced fitness in an environment may indeed persist if its presence is facilitated another species, e.g. decomposition products from one microbial taxa can provide a source of nutrients for another (Tardy et al. 2014). These relationships also impact other aspects of the community, including life history strategies (Barberán et al. 2012) and ecological traits (Williams et al. 2014), and

subsequently have downstream effects on ecosystem functioning and its stability (Tilman 1982, HilleRisLambers et al. 2012). While timescales on the order of days are sufficient to observe compositional changes in bacterioplankton assemblages (Hewson et al. 2006, Redford and Fierer 2009), there is little information on how co-occurrence patterns associated with compositional changes arise and vary along environmental gradients.

Another unresolved question is whether different lineages change over time at different rates (Fierer et al. 2010). Few studies have differentiated among bacterial, archaeal, and fungal community changes despite them differing in their metabolic abilities (Morriën et al. 2017). For example, whether fungi or bacteria are the first to establish in a site may have strong consequences for subsequent species composition. Aquatic saprophytic fungi can decompose some of the most recalcitrant organic compounds from humic substances, subsequently allowing bacteria to colonize humic environments and break down these compounds further if they arrive after the fungi (Grossart and Rojas-Jimenez 2016). Further support for the idea that the arrival of one lineage may influence the rates of change of other lineages is that taxa from different lineages interact metabolically. For example, bacterial growth can be promoted by fungal exudates (Pion et al. 2013, Ponomarova and Patil 2015). While much less is known about archaea, they may vary less than other lineages through time because they respond less to environmental gradients like temperature or organic matter availability (Pala et al. 2018). These examples suggest that how environmental filters influence the initial establishment of different kingdoms will have subsequent consequences for microbial abundances and co-occurrence patterns.

Here we tested how different environmental scales influenced the rate at which microbial community composition changed over time. We also tested whether these changes occurred synchronously across different environments given the same initial

communities. We carried out our study in lake sediment as it allowed us to test the effects of both small- and large-scale environmental conditions. Specifically, we manipulated the small-scale environment by creating sediments with different terrestrial organic matter (t-OM) and placing these in three lakes differing in trophic status (i.e. large-scale environment). We then characterized microbial communities over a two-month period and tested how the rates at which they diverged from the original t-OM were influenced by environmental conditions. We predicted that communities would diverge fastest from the community present in the original t-OM under nutrient-rich conditions, because these provide more opportunities for growth and for colonization by novel species. We then asked whether microbial communities became dissimilar from one another with time and, if so, whether this was primarily due to differences within- or across lakes (i.e. sediment versus overlying waters). As species will first be filtered from the species pool by large-scale environmental conditions (Wiens and Donoghue 2004, de Bello et al. 2013), we predicted that divergence rates would be primarily driven by differences in overlying waters. We also expected different microbial kingdoms to change at different rates because they respond differently to their abiotic and biotic surroundings. Finally, we tested which environmental scale was most associated with temporal changes in species co-occurrence. We predicted that sediment conditions would be the strongest predictor of connectivity because small-scale environmental conditions should act on abundances once communities pass through larger scale filters (de Bello et al. 2013).

4.3 Materials and methods

4.3.1 Study design

We submerged experimental mesocosms in the nearshore region of three small lakes outside of Sudbury, Canada (46°29’24”N, 81°00’36”W) that differed in their overlying

water quality. Two of the lakes were Lake Laurentian (46°27’30”N, 80°56’0”W; area = 1.57 km^2) and Swan Lake (46°21’59”N, 81°3’49”W; 0.06 km^2). Following Williamson et al. (1999), we classified and hereafter refer to these lakes as “mixotrophic” and “oligotrophic”, respectively. Lake Laurentian is mixotrophic as it has mean ± SE total phosphorus (TP) concentrations from summer mid-lake surface grabs of 35.2 ± 2.5 μg L^{-1} and colored dissolved organic carbon (CDOC; i.e. absorption coefficient at 320 nm) concentrations of 26 m^{-1}. By contrast, Swan Lake is oligotrophic because it has much lower TP and CDOC of 9.3 ± 0.4 μg L^{-1} and 1.5 m^{-1}, respectively. Despite these differences, the two lakes are surrounded by similar early-successional forest and experience minimal human disturbance. This description contrasts with our third study site, Ramsey Lake (46°28’42”N 80°56’30”W; 7.96 km^2), which has an extensively urbanized shoreline. We refer to Ramsey as a “mesotrophic” lake after Vollenweider and Kerekes (1982) as its TP and CDOC values were halfway between the oligotrophic and eutrophic statuses described by Williamson et al. (1999), with values of 8.2 ± 0.5 μg L^{-1} and 9.2 m^{-1}, respectively.

Mesocosms with different types of t-OM were placed on the nearshore bottom of each lake beneath 0.30-0.75 m of water during July 2015. The mesocosms were constructed out of HDPE containers that measured 50.8 cm × 38.1 cm × 12.7 cm and were filled with ca. 15 L of sediment after Tanentzap et al. (2017). Briefly, we added 5, 25, and 50% t-OM on a dry-weight basis to 7 kg of locally sourced inorganic material with particle sizes and vertical structuring of all material mimicking natural lake sediments (Tanentzap et al. 2017). For each t-OM quantity, material was comprised of either primarily deciduous, coniferous, or mixed litterfall collected from nearby forests. The deciduous treatment contained 66% litterfall by dry-weight mainly from *Acer rubrum*, *Betula papyrifera*, *Populus tremuloides*, and *Quercus* spp. Coniferous litterfall dominated by *Pinus resinosa* comprised the remaining material. In the coniferous

treatment, the ratios were reversed, with *Pinus* representing 66% of dry-mass and deciduous litter the remainder. The mixed treatment had equal dry-masses of both deciduous and coniferous material. Each treatment was then replicated three times, resulting in a total of 3 t-OM quantities × 3 t-OM qualities × 3 replicates (total n = 27 per lake). Mesocosms were arranged in a block design between two sampling bays, submerged in rows, and covered with a 1 mm × 1 mm nylon mesh screen to standardize the percentage of sunlight reaching the sediment surface. After one month, we made an 8 cm slit in the center of each screen to collect sediment. Importantly, sediment samples taken from our mesocosms reflect the biogeochemistry of the surrounding natural lake sediment (Tanentzap et al. 2017), allowing us to extrapolate our findings to field conditions.

4.3.2 Microbial community characterization

Microbial communities were characterized from surface sediment grabs (ca. top 5-cm) collected from all mesocosms during three sampling periods in 2015: 10 to 12 August, 7 to 9 September, and 5 to 7 October (total n = 81 mesocosms × 3 dates = 243). 16S and ITS primers targeting archaea, bacteria, and fungi were used to construct sequencing libraries after all environmental DNA was extracted (see Appendix C.1 for further details). Libraries were then sequenced on an Illumina MiSeq (600 cycles, paired-end) at an average (±SE) read depth of 15,781 (± 514) and 26,573 (± 1,879) reads for the 16S and ITS sequencing runs. DNA was also extracted in triplicate from the original deciduous and coniferous t-OM and sequenced alongside the other samples (Appendix C.1).

We inferred amplicon sequence variants (ASVs) present in each sample from the raw sequences and taxonomically-annotated them using DADA2 (Callahan et al. 2016a; Appendix C.1). Unlike operational taxonomic units (OTUs), which are arbitrarily

determined by a 97% threshold sequence similarity, each ASV is defined as a unique sequence thanks to a method that controls for amplification and sequencing errors (Callahan et al. 2016a). ASVs therefore represent unique microbial taxa, and yield more accurate and reproducible amplicon data across studies than OTUs (Callahan et al. 2017). From all samples, we removed 46 ASVs out of 42,668 that were present in three negative water-only controls with a relative abundance of >1%. As we were interested in tracking how the bulk of the communities changed with time, we chose to focus our analysis on the 1% most abundant taxa, which together comprised 996 taxa and on average (±SE) 72% (± 1%) of the total reads in each sample. We controlled for the differences in the number of reads per sample due to sequencing biases with a widely-applied variance stabilizing transformation (Dillies et al. 2013, McMurdie and Holmes 2014, Weiss et al. 2017) using the R package DESeq2 (Love et al. 2014). All downstream analyses were performed on the DESeq-transformed data. The raw sequences were deposited in EBI under the project accession number ERP110084.

4.3.3 Do microbial communities diverge faster from starting leaf material because of differences in sediment or lake conditions?

We first assessed how quickly microbial communities diverged from the original mesocosm t-OM, i.e. leaf material. We visualized all the mesocosms with a non-metric multidimensional scaling (NMDS) ordination with Bray-Curtis distances, and then measured the Euclidian distance between the NMDS scores of each mesocosm and the mean centroid of the original leaf material. To test whether the rate at which communities diverged from the original leaf material varied among lakes, we fitted a linear model to the distances and included sampling day, lake, and an interaction term between sampling day and lake as predictors. We fitted a separate model for each kingdom.

We also tested whether compositional differences were associated with sediment porewater conditions within each lake. We extracted 45 mL of porewater on each sampling occasion and immediately measured pH with a handheld meter (HI 9126, Hanna instruments, Woonsocket, RI, USA). We then filtered 25 mL of each sample through a 0.5 μm glass fibre filter (Macherey-Nagel MN 85/90) and into a 20-mL glass scintillation vial, which was pre-acidified for a pH of approximately 2-3 to avoid the effects of metal quenching of DOM fluorescence (Spencer et al. 2007). In the lab, we measured two widely used DOM metrics using a Cary 60 UV Vis spectrophotometer and a Cary Eclipse fluorescent spectrophotometer (Agilent Technologies, Santa Clara, CA, USA). The first DOM metric was the specific UV254 absorbance (SUVA), which is an index of the average aromatic fraction of DOM per unit DOC, itself measured on a Shimadzu TOC-5000A (Shimadzu Co, Columbia, MD, USA). Higher SUVA values indicate higher molecular weight DOM that tends to be more difficult for microbes to break down (Sinsabaugh et al. 1997, Lavonen et al. 2015). We also corrected SUVA values for iron, which absorbs UV at a similar wavelength to SUVA and can artificially inflate SUVA measurements (Weishaar et al. 2003, O'Donnell et al. 2012). Total iron concentrations were measured using the FerroVer method (Hach Company 2014) on a Hach DR3900 spectrophotometer (HACH, Loveland, CO, USA). The second DOM metric was the humification index (HIX), for which higher values indicate less structurally complex DOM and increased humic substance content (Fellman et al. 2010). To achieve normality, SUVA, DOC, and HIX measurements were log-transformed. We used the "envfit" function in the "vegan" R package (Oksanen et al. 2013) to correlate pH, SUVA, DOC, and HIX with the NMDS scores and determined significance using 999 permutations.

4.3.4 Do microbial communities become dissimilar faster because of differences in sediment or lake conditions?

We then tested whether the rate at which microbial communities became dissimilar with time differed with sediment conditions and among lakes. We calculated the dissimilarity between mesocosms that received identical t-OM treatments and were in the same position in our experimental block design but located in different lakes. For a single mesocosm in the block design, we could therefore have three possible comparisons. We used the Bray-Curtis dissimilarity index for our calculations, for which a value of 0 indicated communities were entirely similar while a value of 1 indicated communities were entirely different, i.e. no overlapping taxa (Bloom 1981). Due to the highly variable size of the ITS fragment among fungi, using phylogenetically-informed distances like Unifrac was not recommended (Adams et al. 2014) and so we did not generate phylogenetic distances between ASVs as we would not have been able to recover them for all kingdoms.

We used a linear model to model how dissimilarity changed over time with different quantities and qualities of t-OM and in different lake comparisons. We added interaction terms between the sampling day and each of t-OM quantity, quality, and lake comparison along with all of their main effects. Sampling bay was also included to account for the blocking design of our experiment. To assess the significance of t-OM quantity, t-OM quality, and the identity of the lake comparison, we separately removed each of the main effects and their interactions and compared the reduced and full models with an ANOVA. Considering the variance of the treatments was not constant, we corrected the standard errors as in Cleasby and Nakagawa (2011) with the “coeftest” function in the “lmtest” R package (Zeileis and Hothorn 2002). Finally, we fitted separate models for each kingdom (i.e. archaea, bacteria, and fungi) to test whether different microbes changed differently over time.

4.3.5 Do microbial taxa form different co-occurrence networks with time because of sediment or lake conditions?

We asked whether individual taxa occurred more with other taxa because of sediment or overlying lake conditions. First, we created separate co-occurrence networks at the start (August) and end (October) of the experiment for each t-OM quantity, t-OM quality, and lake using the "igraph" R package (Csárdi and Nepusz 2006). A total of 18 networks were therefore generated, from which we estimated the number of degrees (i.e. co-occurrence among each taxa, or connectivity; Tylianakis et al. 2010) for each mesocosm. Most taxa (90%) were connected to all the other taxa (i.e. >900 degrees per individual taxa). Therefore, we only constructed networks for the 10% most abundant taxa (n = 100) because there was much more variation in connectivity (i.e. the coefficient of variation for 100 taxa was about 20% higher than for 996 taxa). We then measured the difference in the number of degrees for each taxa between the end and start of the experiment for each t-OM quantity, t-OM quality, and lake. We assessed how the number of degrees per mesocosm changed over time using a linear model that included each level of either t-OM quantity, t-OM quality, or lake as a predictor, and allowed these effects to vary among kingdoms.

Where changes in connectivity were statistically significant, we also measured how the number of mesocosms in which each pairwise combination of taxa co-occurred changed with time. We subset our data to the 10 most abundant taxa across all samples for visualization purposes. We plotted the connections between these taxa at the start and at the end of the experiment using the "circlize" R package (Gu et al. 2014) and assessed how connectedness changed in the 10 taxa with pairwise *t*-tests.

4.4 Results

4.4.1 Drivers of divergence from original community

We found that microbial communities in the mixotrophic and oligotrophic lakes diverged the fastest from the original leaf community (t_{237} = 3.58, P < 0.001 and t_{237} = 3.41, P < 0.001, respectively; Fig. 4.1; Table C.1). In each lake, sediment communities were driven away from the original leaf community as pH and HIX correspondingly increased (R^2 = 0.05, P = 0.001 and R^2 = 0.04, P = 0.009, respectively). This observation was especially strong in the mixotrophic lake where pH and HIX were higher (Fig. 4.1a, Fig. C.1). SUVA and DOC were not associated with changes in community composition (Table C.2).

Kingdoms differed in their responses. Archaea, which only comprised 4% of all DESeq-transformed reads, diverged a magnitude faster from the original leaf community than bacteria and fungi, which represented 52% and 44% of all reads, respectively. While archaeal and bacterial communities displayed a similar pattern among lakes – and that mirrored the microbial communities' overall response – fungal communities consistently changed in an opposite direction to bacteria (Fig. 4.1b). Surprisingly, in mesotrophic conditions, the bacterial community did not diverge from the original community as expected if the communities became progressively colonized with local microbes. Rather, the community became more like the original leaf material (t_{237} = -5.69, P < 0.001; Fig. 4.1b), suggesting that abiotic conditions and/or biotic interactions in the mesotrophic lake may have favored bacteria present on the original leaf material.

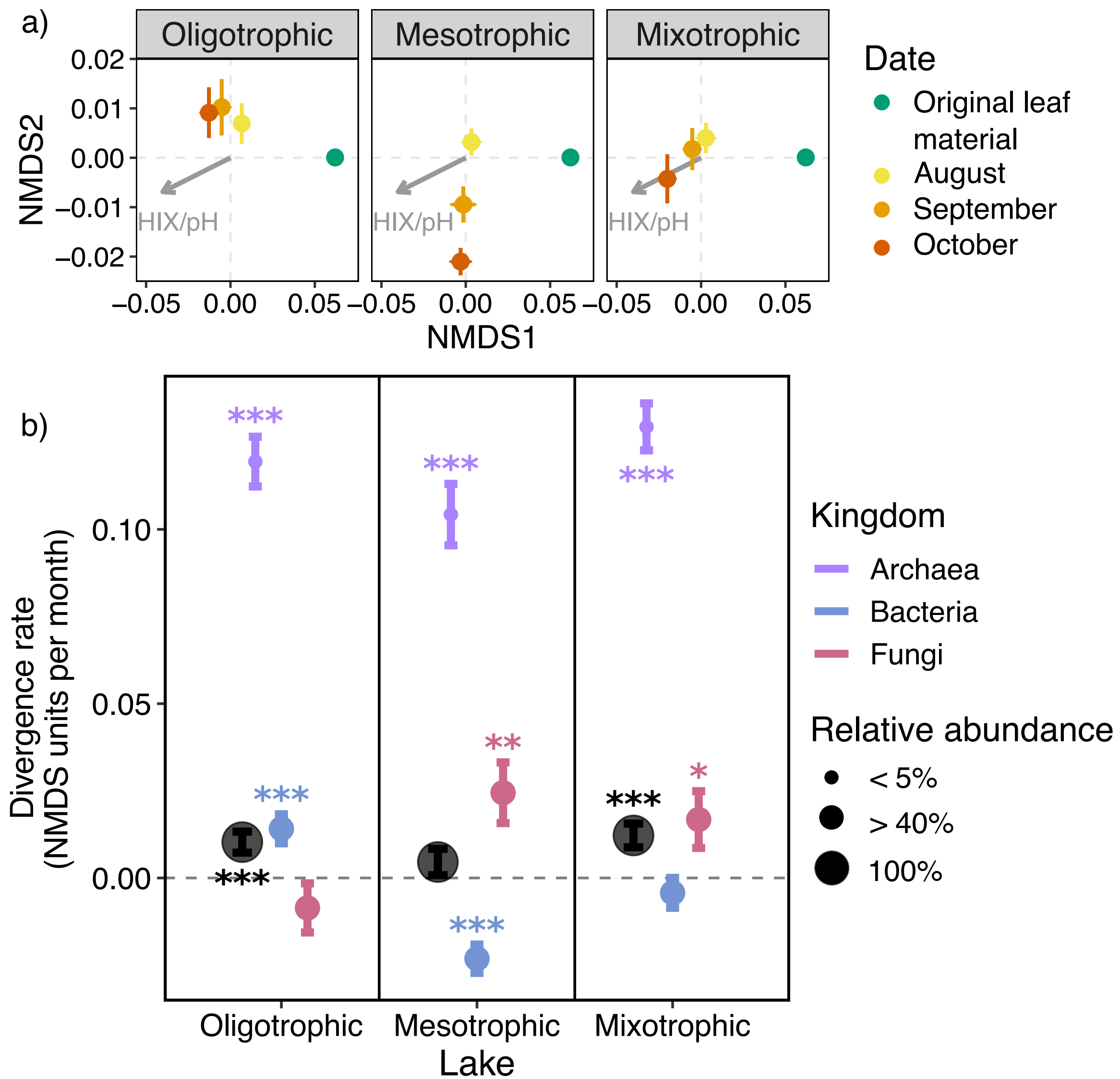

Fig. 4.1 | Microbial communities diverged, on average across kingdoms, from the original leaf community. a) Divergence from the overall microbial community (i.e. across all kingdoms) followed gradients of increasing pH and HIX, especially in the mixotrophic lake. Biplot vectors for pH and HIX were overlapping and appear indistinguishable from each other. b) Divergence rates from the original leaf community were fastest in the oligotrophic and mixotrophic lakes but differed across kingdoms. Black points and lines represent the overall microbial community. Relative abundances were measured as the number of DESeq-transformed reads for each kingdom out of the all DESeq-transformed reads across lakes. Slopes significantly different from 0 denoted by * P = <0.05, ** P = <0.01, *** P = <0.001.

4.4.2 Community similarity across sediment conditions and lakes

We found that microbial communities became more dissimilar with time both because of the sediment and lake conditions (Table C.3). Removing t-OM quantity from the full model subsequently reduced the R^2 from 0.42 to 0.24 as compared with a reduction to 0.35 and 0.39 from removing the effects of lake identity and t-OM quality, respectively (Table C.3).

Communities diverged across all t-OM treatments and did so at a steeper rate with increasing levels of t-OM quantity. Communities in 50% t-OM quantity diverged about twice as fast than those in the 5% t-OM treatment (Fig. 4.2a). For example, communities in the 5% t-OM treatments were 0.46% (95% CI: 0.43-0.49%) more dissimilar at the end of the experiment whereas those in 50% were 1.49% (95% CI: 1.46-1.52%) more dissimilar. Although absolute effects may seem small, they were calculated over two months and involved 996 taxa. Increasing t-OM concentrations from 5% to 25% and from 25% to 50% did not further influence the rate of divergence (Fig. 4.2a; Table C.4).

We also found that communities in coniferous-dominated mesocosms diverged about 1.6 times faster than those in deciduous-dominated t-OM (t_{228} = -2.25, P = 0.025; Fig. 4.2b). Specifically, while dissimilarity was comparable between coniferous and deciduous treatments in August, by October dissimilarity had increased by 1.53% (CI: 1.47-1.60%) in the coniferous treatments compared to a 0.75% (CI: 0.69-0.82%) increase in the deciduous treatments.

Finally, microbial communities in the oligotrophic and mesotrophic lakes diverged over time from those in the mixotrophic lake and did so at similar rates (Fig. 4.3; Table C.4). There was no difference in the divergence rate between microbes in oligotrophic and mesotrophic conditions (Fig. 4.3; Table C.4).

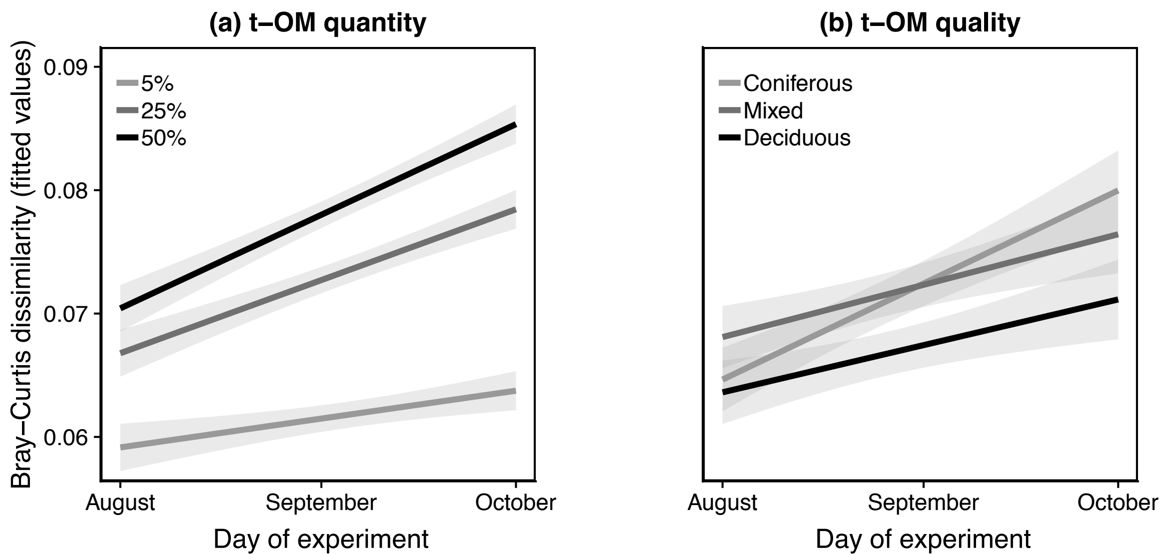

Fig. 4.2 | Communities became increasingly dissimilar with time. Bray-Curtis dissimilarity index increased faster a) at higher t-OM quantity and b) in coniferous-dominated t-OM qualities. Lines are mean model fit at mean values of the other variables. Polygons are 95% confidence intervals.

Microbes belonging to different kingdoms diverged differentially with time (Fig. 4.2). Consistent with our previous findings (Fig. 4.1b), archaeal communities diverged with time across all pairwise lake comparisons at an order of magnitude faster than either bacteria or fungi. Changes in archaeal community composition were accompanied by higher archaeal abundances with time (t_{8503} = 24.77, P<0.001, Table C.5). Similarly, fungal abundances increased with time (t_{8503} = 9.95, P<0.001, Table C.5), but bacterial ones decreased (t_{8503} = -3.37, P<0.01, Table C.5). While the slowest divergence rates for archaea occurred between mesotrophic and mixotrophic lakes, fungi diverged fastest in this pairwise comparison (Fig. 4.3). Fungi in the oligotrophic lake did not change differently from fungi in other trophic statuses (Fig. 4.3). Bacteria, the most prevalent kingdom, changed more slowly than archaea and fungi, significantly diverging only when comparing mixotrophic and oligotrophic communities (Fig. 4.3).

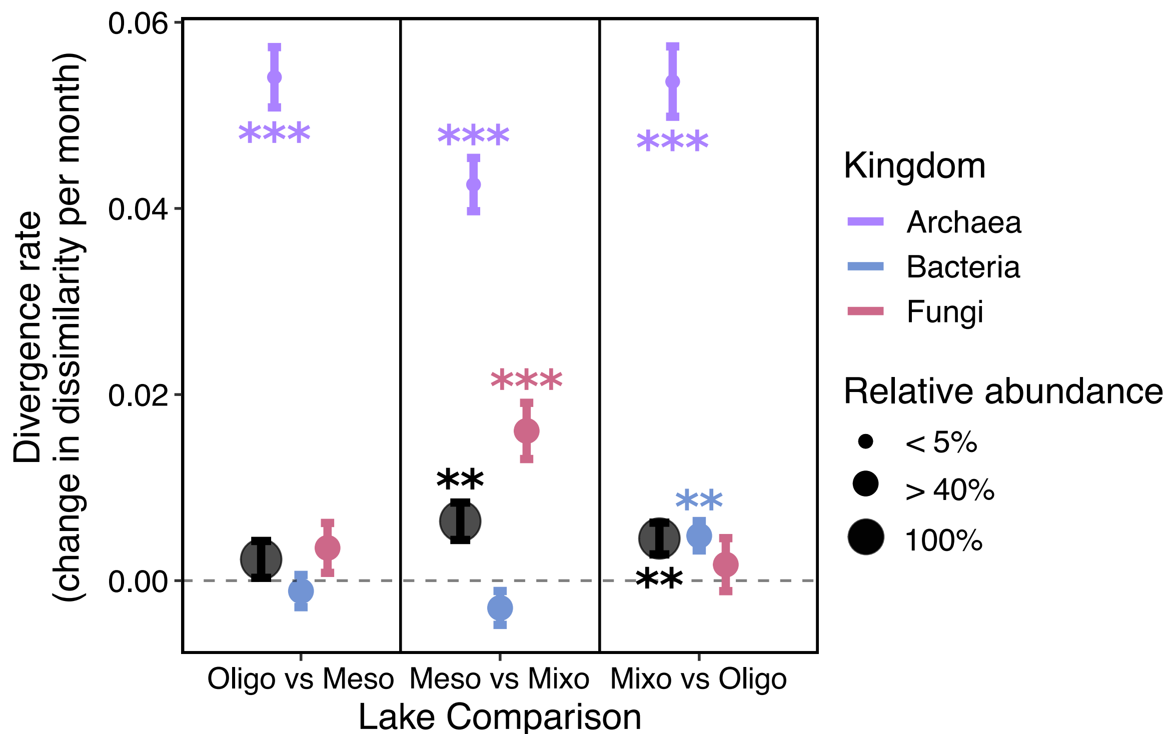

Fig. 4.3 | Comparisons with the mixotrophic lake diverged faster, across kingdoms, than comparisons with other trophic conditions. Points are mean slope ± SE of the monthly pairwise change in similarity between two lakes. Black indicates all microbes (i.e. across all kingdoms). Relative abundances were measured as the number of DESeq-transformed reads for each kingdom out of the all DESeq-transformed reads across lakes. Black points and lines represent the overall microbial community. Slopes significantly different from 0 denoted by ** P = <0.01 and *** P = <0.001.

4.4.3 Drivers of connectivity among taxa

Only t-OM quantity rather than t-OM quality or lake identity influenced how the number of co-occurrences between taxa changed with time. Specifically, while there was no difference in the number of degrees for taxa at 5% t-OM between the start and end of the experiment (Table C.6), taxa present in the 25% and 50% t-OM treatments became more interconnected (Fig. 4.4a). Taxa in these treatments gained on average 6.64 and 7.24 degrees over two months (t_{295} = 2.07, P = 0.039 and t_{295} = 2.26 and P = 0.025, respectively). Connectivity did not vary among kingdoms (Table C.6).

Additionally, we found that the 10 most abundant taxa co-occurred in more mesocosms by the end of the experiment at all treatments of t-OM quantity (pairwise *t*-test: P always < 0.001; Fig. 4.4b; see Table C.7 for taxa names).

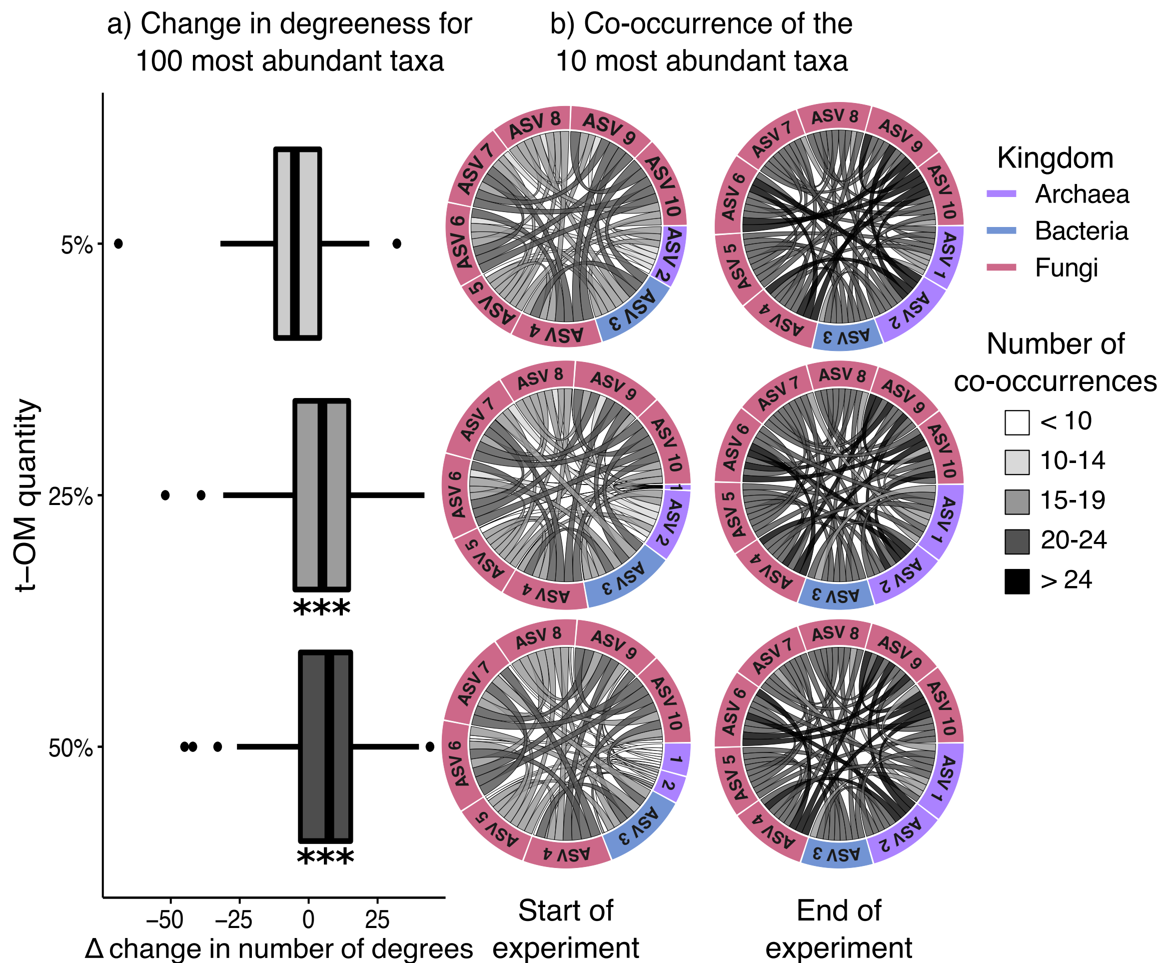

Fig. 4.4 | Microbial taxa co-occurred more over time at higher t-OM quantity. We calculated a) degrees as the number of co-occurrences between each taxa and all others for the 100 most abundant taxa and b) the number of mesocosms in which each pairwise combination of the 10 most abundant taxa co-occurred (see Table C.7 for taxa identity). *** indicates P value < 0.001.

4.5 Discussion

Our results reveal that both small- and large-scale environmental filters, as measured by sediment characteristics and lake trophic status, respectively, influence the rate at which

microbial communities develop. To our knowledge, nothing is known about how temporal changes in lake sediment microbial communities vary at different environmental scales. Here we found that communities became more dissimilar with time despite being derived from the same initial leaf material. We also found that dissimilarities between mesocosms were primarily driven by the quantity of sediment t-OM, which was also the only predictor of the connectivity among taxa during our experiment. These findings implicate small- as opposed to large-scale environmental conditions as more important for filtering microbial community composition, in contrast to findings from plants (Wiens and Donoghue 2004, de Bello et al. 2013, Kraft et al. 2015). Considering microbes typically range from 1-5 μm in size, it is not entirely surprising that they were filtered by much finer environmental conditions than larger organisms. Finally, we provide among the first evidence that microbial kingdoms changed differently with time and across environmental gradients, with archaea changing an order of magnitude faster than bacteria or fungi, which themselves often displayed contrasting responses.

Communities diverged faster within lakes because they possessed higher t-OM concentrations rather than because they were in different lakes residing on the same sediment conditions. Despite its high humic and lignin content, t-OM is bioavailable to microbes (Guillemette and del Giorgio 2011, Lapierre et al. 2013) and is primarily allocated to biomass (Guillemette et al. 2016, Fitch et al. 2018). Consequently, the increase in abundances associated with higher t-OM can result in compositional changes and faster divergence rates, especially in the early stages of community development (Fierer et al. 2010). Consistent with this interpretation, t-OM has been shown to enhance fluctuations in individual species' abundances (Chapter 3) and to filter aquatic bacterial community composition (Muscarella et al. 2016). Increasing the quantity of t-OM also increased connectivity. Thus, in addition to accelerating divergence rates

between communities, environmental filtering simultaneously strengthened networks within these communities, thereby reinforcing the differences between communities. Our results are important considering that metabolic dependencies that form between taxa (Zelezniak et al. 2015) may sustain these differences and lead to long-lasting co-occurrence patterns with consequences for community stability (HilleRisLambers et al. 2012).

Additionally, we found that communities in coniferous-dominated sediments diverged the fastest, further implicating small-scale environmental conditions as the primary filter over microbial community composition. Coniferous litter generally contains fewer phenolic compounds than deciduous litter (Kuiters and Sarink 1986, Emilson et al. 2018). As phenolics inhibit extracellular enzymes involved in decomposition (Wetzel 1992), they can restrict species establishment and thus turnover. Their lower abundance in the coniferous sediments may have consequently promoted more abundant and diverse microbial assemblages (Yakimovich et al. 2018). Our findings therefore suggest that rates of microbial community change may become reduced with expected northward shifts in deciduous tree species that will modify the composition of the t-OM delivered to inland waters (Boisvert-Marsh et al. 2014, McKenney et al. 2014).

Communities changed faster in mixotrophic conditions, which may have more favorable conditions for microbial growth than nutrient-limited conditions where microbes instead metabolically cycle carbon (Hessen 1992, Reche et al. 1998, Zwart et al. 2015). Microbes can similarly allocate larger supplies of carbon towards biomass in darker (i.e. mixotrophic) lakes as compared to lakes with more light (Fitch et al. 2018), wherein high photo-oxidation modifies t-OM structure and makes it less bioavailable (Kirchman 2013, Ward and Cory 2016). These ideas may equally explain why temporal changes in community composition more closely followed the HIX gradient in the

mixotrophic lake, which was larger than elsewhere (Fig. C.1b) and is usually associated with variation in nutrients and photo-oxidation (Lindstrom 2000, Schallenberg et al. 2017). Additionally, we found that the overall microbial community diverged from the initial leaf litter in the mixotrophic and oligotrophic lakes and did so faster with higher pH and more humic material across all lakes. These results align with previous studies demonstrating the limited ability of terrestrial taxa to grow and establish in aquatic ecosystems (Ruiz-Gonzalez et al. 2015b) and show that both sediment pH and humification are major drivers of community assembly and composition (Ruiz-Gonzalez et al. 2015a, Amaral et al. 2016, Fitch et al. 2018, Tripathi et al. 2018). Importantly, these findings suggest that, in addition to changes in forest cover that will modify lake sediment composition, rates of microbial community change will also be influenced by the browning of northern lake waters due to climate warming and increased runoff (Solomon et al. 2015, Finstad et al. 2016).

Kingdoms varied by up to 10-times in how quickly they changed, and as they are dominated by different traits, such variation could have far-reaching consequences for ecosystem functioning. On average, communities of archaea – the only organisms able to produce the potent greenhouse gas methane (Garcia et al. 2000) – changed ten times faster than fungi and bacteria regardless of the environmental conditions. Considering that archaea only represented 4% of all the taxa in our analyses, we were surprised to find that two archaeal methanogens were the 5th and 10th most abundant taxa in our sediments. Lakes release 6-16% of global methane emissions (Borrel et al. 2011) so understanding why these communities changed so rapidly is important for predictions of whole-lake C cycling. One reason for these particularly rapid changes may be that archaea were in much lower abundances on the starting material, resulting in less intra-specific competition. Individual archaeal species could therefore grow rapidly, as reflected by their increased abundance with time, and subsequently drive

microbial community changes through time (Fierer et al. 2010). Additionally, bacteria and fungi tended to change in opposing ways, consistent with findings across a range of habitats (Bahram et al. 2018, Yakimovich et al. 2018). Such inter-kingdom antagonism has been attributed to fungi's high competitive ability and antibiotic production (Bahram et al. 2018). Fungi can use the complex carbon substrates that plant cell walls are made of better than bacteria, especially in high nutrient conditions (Koranda et al. 2014). Fungi's competitive advantage over bacteria may therefore explain why they grew and diverged faster in the mesotrophic and mixotrophic conditions, where nutrient levels were higher. Finally, arrival order may also help explain these differences. Lake fungal communities are partly comprised of terrestrial fungi from imported leaves (Barlocher and Boddy 2016) and these taxa usually dominate the initial stages of plant litter decomposition (Kuehn 2016). It is therefore likely that fungi that were better adapted than bacteria to breaking down these t-OM were already present on the original leaf material, allowing them to colonize the sediment first and use its carbon resources, thereby preempting later arrivals from establishing. Such "priority" effects have been found in fungal (Dickie et al. 2012) and bacterial communities (Rummens et al. 2018) but not yet between kingdoms.

Our results are important as understanding how communities change with time and identifying the drivers of these temporal differences will enhance the predictability of community dynamics and ecosystem functioning (Grenfell et al. 1998, Kent et al. 2007). In summary, we found that, given the same starting material, microbial communities did not change synchronously with time. Community composition and species co-occurrence instead differed along small-scale environmental filters. Our study did not account for stochastic processes that randomly modify the species pool and can thereby influence community change (Vellend 2010). However, deterministic processes related to environmental filtering are thought to be more important for

microbes, suggesting that the effects of any stochastic process were likely minor (Hanson et al. 2012, Lindström and Langenheder 2012, Nermegut et al. 2013, Bahram et al. 2018). Predicting how communities change with time and their scale dependence is particularly valuable in regions where small- and large-scale changes are anticipated, such as in boreal ecosystems where warmer climates and increased t-OM export will impact lakes and rivers (Creed et al. 2018). Here we have shown that the outcomes of these changes in microbial communities will vary depending on the amount and type of t-OM delivered to inland waters but also with lake-specific characteristics. Any future changes to lake waters or sediments are therefore likely to have far-reaching consequences for key ecosystem processes like carbon cycling by modifying lake sediment microbial communities.

Chapter 5

General discussion

Lake sediment microbes are at the heart of ecosystem functioning in freshwater systems: by processing terrestrial organic matter (t-OM), they fuel both food web production and carbon cycling. The aim of this thesis was to advance our current understanding of microbial community assembly and functioning along terrestrial resource gradients in boreal lake sediments. I have shown that positive biodiversity-ecosystem functioning relationships occur in complex microbial communities and, importantly, that these relationships depend directly on present-day environmental conditions and indirectly on historical legacies (**Chapter 2**). In addition to enhancing CO_2 production, I have also shown that more diverse microbial communities can stabilize this ecosystem function through time by reducing its temporal variation, and this happens regardless of the environmental gradients (**Chapter 3**). Nevertheless, my results show that microbial abundances are stabilized by higher t-OM inputs rather than by diversity levels, and this stabilizing effect is the result of increased community-level species asynchrony (**Chapter 3**). Lastly, my work reveals the importance of small-scale environmental filters such as sediment conditions for determining the rates of change of microbial communities and the strength of their co-occurrence networks in their early development, which will subsequently impact lake sediment ecosystem functioning (**Chapter 4**).

In **Chapter 1**, I laid out three questions that this thesis aimed to answer:

1) How do temporal processes influence aquatic microbial community assembly and structure (**Chapters 2** and **4**)?
2) How do spatial environmental gradients related to terrestrial inputs influence community structure (**Chapters 2-4**)?
3) How do the environment and community structure individually and interactively influence ecosystem function (**Chapters 2** and **3**)?

In this final chapter, I discuss how the analyses presented in this thesis have addressed these questions. I also critically evaluate the limitations of my work and identify possible avenues for future research that have emerged in the light of my results.

5.1 Past historical events directly influence community structure and small-scale conditions drive community changes through time

There is now ample evidence for the role of past and present-day events in shaping microbial communities (Martiny et al. 2006, Andam et al. 2016, Fierer 2017, Martiny et al. 2017, Delgado-Baquerizo et al. 2018b). Yet, how much of community structure is explained by either of these ecosystem properties in natural systems is still largely unexplored. In this thesis, I found that the similarity of bacterial composition between pair-wise samples decreased with their geographic distance, consistent with studies that highlight the effect of past history and biogeography on community structure (Zinger et al. 2014, Powell et al. 2015, Andam et al. 2016). I also showed that site-level differences in bacterial communities influenced multiple aspects of community structure (i.e. normalized bacterial abundance, taxonomic diversity, and normalized oxidase gene abundance) while the present-day environment did not. These site-level differences

directly explained up to 20% of community structure but they also indirectly explained additional variation because the measures of community structure influenced one another. For example, bacterial abundances explained up to 34% of the variation in taxonomic diversity. These findings are particularly important as they suggest that past historical events rather than present-day habitat-based filtering indirectly influenced ecosystem functioning by modifying community structure (Hendershot et al. 2017, Martiny et al. 2017). Examples of such past events that shape microbial biogeography include colonization, diversification, extinction, and dispersal limitation (Martiny et al. 2006). In my study however, these site-level differences could not be attributed to dispersal-limitation as most taxa were found across all sites, evidence that abundances rather than presence varied across space.

Following on from these results, the next logical step was to understand how past historical events, such as colonization dynamics, could have influenced community structure in lake sediment microbes. To do so, I looked at how environmental filters at different spatial scales influenced changes in microbial communities during the early stages of their development. Research on microbial succession has been limited to soils, biofilms, and glaciers (Jackson et al. 2001, Martiny et al. 2003, Nemergut et al. 2007, Schütte et al. 2010, Dini-Andreote et al. 2015, Smith et al. 2015) and has rarely focused on the early stages of succession despite these stages being essential for subsequent assembly and functioning (Fierer et al. 2010). I found that small-scale abiotic conditions primarily drove changes in community composition through time. Although this result is not entirely aligned with what is know from vegetation studies, where large-scale filters usually filter species composition (Wiens and Donoghue 2004, de Bello et al. 2013), it is not surprising considering the much finer spatial scale of microbes compared to plants. My results however concur with the fact that small- rather than large-scale environmental filters influence the formation of species co-occurrence patterns (Diaz et

al. 1998, Chesson 2000, de Bello et al. 2013). This small-scale filtering of connectivity among taxa was due to variation in the amount of t-OM in the sediment, but other small-scale filters like t-OM quality and larger-scale lake conditions also mattered. Together, these results suggest that while microbes may be found across wide environmental ranges or originate from the same material, they may not flourish everywhere. These differences in composition and abundances likely occur because microbes respond strongly to both the sediment and the overlying water conditions (Ruiz-Gonzalez et al. 2015a, Nino-Garcia et al. 2016).

Ultimately, allochthonous inputs largely influenced the processes of assembly described above, possibly affecting the current composition of lake sediment microbiomes. Compositional legacies have been shown to persist for up to 3 years in fungal communities (Martiny et al. 2017). So while present-day environmental filters like OM quantity did not directly influence patterns of bacterial abundance or diversity in **Chapter 2**, it is likely that past processes of colonization were influenced by t-OM quantity as in **Chapter 4**. This past filtering would have structured the community at the time of colonization and resulted in the distinct assemblages that persisted through time differently from neighboring communities.

5.2 t-OM quantity and quality are the primary spatial influences over microbial community structure

Variations in the amount and type of OM also drove spatial patterns of microbial community structure. In **Chapter 2**, community similarity increased with environmental similarity across eight sites, suggesting that present-day environments influenced microbial community composition. More precisely, community composition was driven both by the amount of terrestrial carbon and by the C:N ratio that are characteristic of OM origin and composition, respectively. In **Chapter 3**, more t-OM

increased community-wide asynchrony: t-OM thus influenced how individual taxa differentially fluctuated in their abundances over time. Finally, in **Chapter 4**, both t-OM quantity and quality influenced how communities changed during their early development, but t-OM quantity explained more of this variation than t-OM quality.

Revealing the importance of allochthonous inputs – particularly t-OM quantity – in structuring lake sediment microbiome improves our understanding of terrestrial resource use in aquatic systems. The bioavailability of t-OM for microbes has long been debated, with some arguing that it is too recalcitrant to be incorporated into biomass (Sollins et al. 1996, Williamson and Morris 1999, Zwart et al. 2016) and others arguing that it is an important carbon resource for microbes (Judd et al. 2006, Lapierre et al. 2013, Jones and Lennon 2015, Guillemette et al. 2016, Fitch et al. 2018). In this thesis, I found that t-OM quantity was a key predictor of microbial community structure, suggesting that allochthonous inputs may be used towards microbial growth and subsequently promote community change. Overall, these results are consistent with others finding that t-OM quantity rather than quality is the primary spatial influence over microbial community structure (Judd et al. 2006, Crump et al. 2012, Besemer et al. 2013, Ruiz-Gonzalez et al. 2015a). Recent work by Ruiz-Gonzalez et al. (2018) showed that terrestrial inputs enhanced the growth and activity of certain aquatic taxa, but that different forest soils did not cause predictable changes in lake microbial communities. They suggested that the different growth patterns between different soil treatments may be linked to priority effects or to changes in the overlying waters that favor certain taxa over others, in line with findings from **Chapter 3**.

5.3 The environment and community structure influence ecosystem functioning interactively but stabilize it independently

While a large number of studies have shown that plant diversity enhances and stabilizes ecosystem functioning (Hooper et al. 2012, Tilman et al. 2012, Hautier et al. 2014, Jucker et al. 2014), evidence for these patterns is equivocal for microbes. In **Chapter 2**, I found a positive B-EF relationship in bacteria, with taxonomic diversity driving 13% of the observed variation in CO_2 production in lake sediments. However, functional diversity did not explain ecosystem functioning. The lack of an effect of functional diversity is not entirely surprising considering that multiple taxa are likely to perform the same tasks associated with a ubiquitous function like C utilization (Carlson et al. 2007, Wertz et al. 2007, Delgado-Baquerizo et al. 2016). Nevertheless, another aspect of community structure – normalized oxidase gene abundance – drove another 13% of CO_2 production. This result highlights that focusing on the identity of traits associated with a response of interest provides more valuable information for predicting ecosystem function than simply looking at functional diversity, consistent with the trait-based approach advocated by many groups (Krause et al. 2014, Wieder et al. 2014, Roger et al. 2016, Trivedi et al. 2016). Trait-based approaches are becoming increasingly popular for microbial community studies but there is still a missing gap in identifying which traits – or genes – most regulate which microbial functions, especially for ubiquitous ones like CO_2 production. In this thesis, I found that oxidase associated genes, rather than hydrolase ones, were most important for CO_2 production in lake sediments. Interestingly, this association was even stronger when there was more terrestrial carbon. At high % terrestrial carbon, oxidase genes were even more active in respiring, suggesting that functional genes interacted with terrestrial carbon to promote CO_2 production.

In addition to promoting CO_2 production, I found that species-rich assemblages stabilized sediment functioning over time. These results closely match those commonly found in plant communities and also those of a recent study done on soil microbial communities (Wagg et al. 2018). Unlike my previous results in **Chapter 2** showing that community structure and the environment interacted to promote CO_2 production, there was no interactive effect between the environment and diversity in stabilizing ecosystem function in **Chapter 3**. Bacterial richness instead stabilized CO_2 production regardless of the environment suggesting that CO_2 production is a function performed so widely that it will be observed across a broad range of environments (Carlson et al. 2007, Wertz et al. 2007). In parallel, the environment stabilized microbial abundances regardless of diversity. Specifically, increasing t-OM quantity led to greater fluctuation of taxa, likely because some taxa can use this resource better than others (Findlay 2003, Berggren et al. 2010b). This increase in community-wide species asynchrony subsequently stabilized the community.

5.4 Limitations and avenues for future work

5.4.1 Improving experimental design and sequencing

A key challenge for microbiome studies is to maintain robust scientific practice while integrating new approaches unique to the field (Knight et al. 2018). This challenge has been exacerbated by the elevated cost of sequencing that sometimes appears to compromise classic experimental considerations, such as sample size, number of replicates, and appropriate controls (e.g. extraction and reagent blanks). Indeed, each extra sample sequenced incurs considerable additional costs (i.e. sum of costs of DNA extraction, library preparation, sequencing, bioinformatics processing) that seem to deter researchers from following stringent ecological procedures, as they usually would

in other systems. Yet, meticulous experimental design is essential for the validity of any microbiome study as well as for cross-study comparisons.

The confounding issues associated with microbiome studies are increasingly highlighted in the literature, with more standardized procedures being proposed (Bálint et al. 2017, Hugerth et al 2017, Knight et al. 2018, Pollock et al. 2018). These more recent papers helped me improve experimental design throughout the course of my thesis. For example, I moved from having no technical replicates in **Chapter 2** to having two per biological sample in **Chapters 3** and **4**. Additionally, although I included negative blanks in my first library preparation, I did not sequence these blanks as they appeared to contain no DNA. However, they may still have contained a few rare contaminant sequences that should ideally have been removed. In later runs, I included these blanks during the sequencing and removed any sequence found in high abundance (>1%) from the rest of the community (as per Flores et al. (2012) and Salter et al. (2014)).

Another source of bias comes from the choice of primers for amplicon sequencing, as different primers preferentially select certain taxa. Zumsteg et al. (2012) for instance suggested that they found more *Euryarchaeota* in younger soils than previous studies had because they used different primers. Primer bias also arises during PCR because the number of cycles may influence how many and which sequences are recovered (Clooney et al. 2016, Knight et al. 2018). These considerations may have been particularly important in my analysis on community changes through time (**Chapter 4**), in which I used two primer pairs: one for prokaryotes and one for eukaryotes. The issues associated with primer bias may be solved if we move away from amplicon sequencing to the newer shotgun sequencing approach (as in **Chapters 2** and **3**), which also has the benefit of providing functional information (Clooney et al. 2016, Ranjan et al. 2016, Knight et al. 2018). However, despite becoming increasingly

popular, shotgun sequencing of microbial communities is not as developed as amplicon sequencing and may not recover taxonomic information as effectively (Tessler et al. 2017). The potential for this approach will increase as its methodology improves and becomes more standardized in the coming years.

5.4.2 Increasing reproducibility with bioinformatics pipelines

Identifying the reads obtained by high-throughput sequencing – taxonomically and functionally – is challenging. There are multiple reference databases used for taxonomic alignment (e.g. SILVA, Greengenes, RDP), and picking one over the other may influence which taxa are recovered (Pollock et al. 2018). Another step that can introduce error is the clustering of sequences. Traditionally, sequences are clustered into OTUs at a 97% similarity threshold into single sequences. This threshold can have strong influences on the number of unique taxa recovered: the clustering is done relative to the particular sequences present in the sequencing run and the resulting clusters of OTUs will therefore not be fully comparable across studies. Such clustering may mean that subtle but real biological variations are missed (Callahan et al. 2016a, Knight et al. 2018). This is why more recent methods that use error profiles to resolve sequence data into exact amplicon sequence variants (ASVs), such as DADA2 (which I used in **Chapter 4**; Callahan et al. 2016a) and Deblur (Amir et al. 2017), are better for assigning taxonomy to sequences than classic methods (Callahan et al. 2017).

These improved methods of taxonomic assignment may also allow better recovery of the rare biosphere (Patin et al. 2013, Callahan et al. 2016a), which has been shown to play a non-negligible role in microbial functioning (Campbell et al. 2011, Shade et al. 2014, Lynch and Neufeld 2015, Jousset et al. 2017). In this thesis, I partly ignored rare taxa by removing singletons (**Chapter 3**), sequences not represented more than 5 times across a sequencing run (**Chapter 2**), and by focusing on the 1% most

abundant taxa (**Chapter 4**). These steps were done to avoid the inclusion of spurious reads from contamination (Brown et al. 2015) and to focus on the most abundant taxa in the latter chapter. This filtering may have limited the scope of my analyses and it would be interesting to include data on the rare biosphere in the future using the most recent advances in the field (Jousset et al. 2017).

Ultimately, the results obtained through sequencing are meaningful so long as the analysis is carried out in a consistent manner across samples (Knight et al. 2018, Pollock et al. 2018). This requirement, however, means that cross-study comparisons are difficult to achieve for microbial communities. Projects like the Earth Microbiome Project (EMP) have tried to address this issue by encouraging researchers to follow a standardized experimental protocol and bioinformatics pipeline (Thompson et al. 2017). The latest version of this pipeline in fact recommends using ASVs rather than OTUs (Thompson et al. 2017). It is likely that as ASVs become more widely used, comparability among studies will be improved. While initiatives like the EMP are admirable, their usage is still confined to a limited number of studies and may be difficult to achieve in the long-run. One reason for limited uptake is that as sequencing protocols and technologies improve at a fast-pace, labs may wish to align themselves to the most recent developments rather than to the slightly older methods used in standardized protocols. Sequencing costs may also limit certain groups' ability to choose specific reagents and kits. Until using standardized protocols becomes the norm – if it ever does – researchers should precisely report how they run their analyses and deposit their raw data and metadata in public repositories like EMBL or MG-RAST (Knight et al. 2018).

5.4.3 Advancing research on microbiome functioning

In the past decade, thanks to the development of sequencing technologies, the field of microbial ecology made tremendous leaps in understanding how microbial communities function (Antwis et al. 2017). This progress is recent though, meaning that some ecological questions that have been thoroughly studied for decades in other systems are still unresolved or debated in microbes. For example, there is conflicting evidence for the B-EF theory in microbes, with some finding positive relationships (Delgado-Baquerizo et al. 2016, Laforest-Lapointe et al. 2017) and others not (Jiang 2007, Becker et al. 2012). In my thesis, I found support for positive B-EF relationships in natural lake sediment communities (**Chapter 2**) but not in artificial sediment communities (**Chapter 3**). It is worth noting that the latter was calculated with OTU richness averaged over a year rather than with a single time-point measure of Shannon's diversity index, which accounts for evenness and abundances of OTUs. Importantly though, the results from my thesis (**Chapters 2-4**) highlight that these B-EF relationships may depend on specific environmental conditions or historical events that structure communities in their early development and through time. A final point is that our understanding of microbial evolutionary processes is still limited and may influence how we measure microbial biodiversity (Antwis et al. 2017). High levels of horizontal gene transfer (Doolittle 1999), large numbers of unidentified microbes, and difficulties associated with delineating microbial "species" (Freudenstein et al. 2016) may skew taxonomic assignments, changing the strength and direction of B-EF relationships across studies.

In addition to improving our knowledge of microbial diversity, it is necessary to further our understanding of ecosystem functioning, particularly of microbial-regulated processes. More functions, like methane production, microbial biomass, and nutrient cycling, should be considered in future B-EF and D-S studies. My results support the expectation that ecosystem function and its stability are influenced by both past- and

present-day environmental conditions (Reed and Martiny 2007, Chase 2010, Peralta et al. 2016) and by various aspects of community structure, including taxonomic diversity (Delgado-Baquerizo et al. 2016, Laforest-Lapointe et al. 2017) and the abundance of specific functional genes (Krause et al. 2014). Disentangling and quantifying these influences further will help improve predictions of biogeochemical cycles that are largely regulated by microbes. Another next step from my thesis that would strengthen our understanding of lake sediment microbiomes would be to link their co-occurrence patterns to functioning and stability (Kara et al. 2013), and to examine these processes over longer periods of time. Finally, my findings reveal that each microbial kingdom is likely to influence functioning differentially, by changing differently through time both in terms of direction and rate (up to an order of magnitude faster for archaea; **Chapter 4**). These findings are particularly relevant as recent work by De Vries et al. (2018) showed that fungal networks were more stable in the face of drought than bacterial ones. Investigating how each microbial kingdom's response to environmental change will affect ecosystem functioning and its stability will be critical in the future.

5.5 Concluding remarks

A missing gap in microbial ecology is how temporal and spatial variation in microbial community structure relates to key environmental processes and geochemical cycles (Antwis et al. 2017). Understanding these links are especially relevant in functionally important systems that are expected to change in the near-future, such as boreal lakes where warmer climates and increased loadings of t-OM will impact sediments that fuel freshwater food webs and carbon cycling (Creed et al. 2018). In this thesis, I have shown that these t-OM inputs drive microbial community change, and also directly and indirectly influence ecosystem functioning and its stability in lake sediments. In the context of a changing planet, my results show that models integrating biodiversity and

trait-based approaches will better predict the outcomes of future environmental changes (Krause et al. 2014, Martiny et al. 2015, Roger et al. 2016, Trivedi et al. 2016). These predictions will also be improved by considering the legacies of microbial communities (Martiny et al. 2017) and by understanding why the important functions undertaken by freshwater microbes can be maintained through time (Downing et al. 2014). Importantly, any future changes to lake waters or sediments that modify microbial communities are likely to have far-reaching consequences for key ecosystem processes like carbon cycling.

Appendix A

Biodiversity-ecosystem functioning relationships in lake sediment microbiomes

A.1 Supplementary methods

A.1.1 Site details

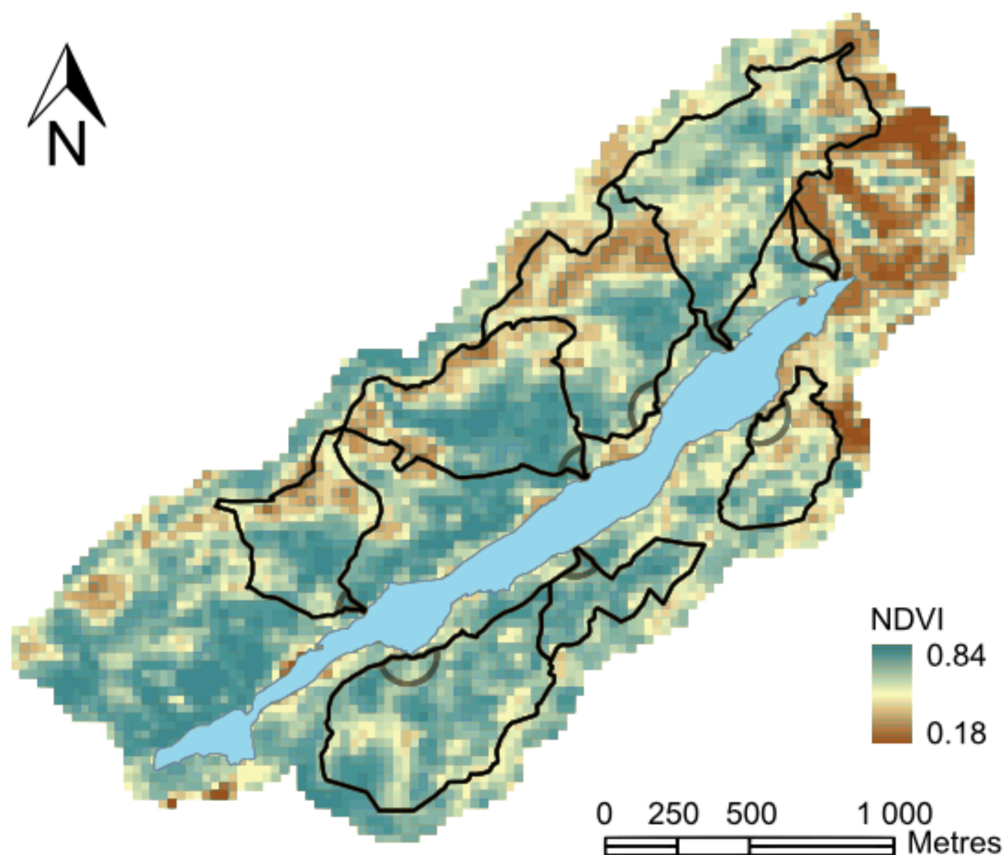

Fig. A.1 | Map of the study sites, showing vegetation density. Normalized difference vegetation index (NDVI, 30m x 30m pixels) within boundaries of study catchments (black line) and riparian zones 100m from stream-lake interface (grey curves) (adapted from Tanentzap et al. 2014).

A.1.2 Measurement of sediment OM quality

In a separate study, we collected 6 surface sediment cores (maximum ~10 cm deep, 5 cm diameter) at each site using either a Wildco hand corer or, where sediments were impenetrable by hand corer, PVC tubes were driven into the sediments with percussion. We collected the cores at approximately 2, 4 and 6 m from shore on each of two longitudinal transects, moving the transect where necessary to avoid boulders. We extruded the cores in the lab into 1 cm increments, and two surficial increments per core (upper 1 and 3 cm) were retained where possible for isotopic analysis (n = 11-12 samples per site, total n = 94).

We then collected leaf litter and surficial soils to characterise the main terrestrial sources contributing to sediment OM. Leaf litter (excluding anything <1 mm) was collected within 5 replicate $0.25m^2$ plots located at random bearings at a distance of 1, 3, 5, 7 and 9 m from the mouth of the drainage streams. We also collected surficial soil samples (excluding anything >1 mm) at random bearings with a hand corer, but collected 6 replicates along two transects at distances of 1, 5 and 9 m from the mouth of the drainage streams.

Finally, to characterise aquatic sources that would contribute to sediment OM in the benthic sites, we collected periphyton and phytoplankton. Periphyton is a mixture of bacteria, phytoplankton, and detritus, which integrates temporal variability in within-lake production (Tanentzap et al. 2014). We collected periphyton by anchoring 6 clay discs (~12 cm diameter) 0.10 m above the sediment in each site, placed out of direct stream discharge areas to avoid accumulation of terrestrial detritus. The discs accumulated growth between 15 August and 5 September 2014, and all material was then carefully scraped into a sample container. We also collected phytoplankton by filtering 3 replicate 12 L water samples from above the sediments on 3 separate dates in early autumn 2014 through 0.2 μm filters and retaining the material on the filters. DOC

concentrations above the delta sites were relatively low (mean ± standard deviation = 2.5 ± 0.29 mg L^{-1}), so the filtrate, which weighed between 2.5-10 mg, would have been primarily derived from phytoplankton.

All samples were analyzed for $\delta^{13}C$, $\delta^{15}N$ and $\delta^{34}S$ using an Isoprime stable isotope ratio mass spectrometer (IRMS) (GV Instruments, Manchester, UK) at the Natural Environment Research Council Life Sciences Mass Spectrometry Facility (East Kilbride and Lancaster, UK). To remove carbonates, sediment and soil samples were acidified with 10% HCl and rinsed in Milli-Q water prior to $\delta^{13}C$ analysis. We also measured the %C, %N, and %S of each sample using the IRMS.

A.1.3 Estimation of isotopic mixing model

We estimated the relative contribution of terrestrial OM sources to sediment by parameterising using a Bayesian three-isotope mixing model (Solomon et al. 2011). The model estimated the proportional contribution of terrestrial soil (ϕ_{SL}) and leaf litter (ϕ_{LF}), and lake periphyton (ϕ_{PR}) and phytoplankton (ϕ_{PY}) organic matter in the surficial sediment of each sample *i* at each site *j*. We let the ratios of each of the l = 3 isotopes ($\delta^{13}C$, $\delta^{15}N$, and $\delta^{34}S$) be described by a linear mixture of the isotope ratios in each of the four sources: soil (SL), leaf litter (LF), periphyton (PR), and phytoplankton (PY). To estimate the relative contributions of each source to the sediments, isotopic ratios in sediment were sampled from a multivariate normal distribution with a mean vector μ_{ij} and $l \times l$ matrix Σ_{ij}:

$$\mu_{ij} = \Phi_{\text{SL}ij}\text{SL}_{ij} + \Phi_{\text{LF}ij}\text{LF}_{ij} + \Phi_{\text{PR}ij}\text{PR}_{ij} + \Phi_{\text{PY}ij}\text{PY}_{ij}$$

$$\Phi_{\text{SL}} + \Phi_{\text{LF}} + \Phi_{\text{PR}} + \Phi_{\text{PY}} = 1. \qquad \text{(eq. A.1)}$$

We propagated the uncertainty into Σ_{ij} by summing the product of the sources and their observed variances, and added this to an estimated l length vector of residual errors ε, with each element sampled from a relatively uninformative uniform prior

between 0 and 30. We then multiplied this vector of standard deviations with an estimated $l \times l$ correlation matrix Ω to derive Σ_{ij}. The prior for Ω was sampled from a LKJ distribution that placed almost uniform support over the estimated correlations (shape parameter $\eta = 2$), with the density slightly more concentrated around the identity matrix (Lewandowski et al. 2009). We accounted for fractionation of δ^{13}C and reduction of δ^{34}S by estimating a correction factor Δ_i from a uniform prior between 0 and 30 and -50 and 0 for δ^{34}S, respectively. We expected a positive correction factor (i.e. enrichment) for the δ^{13}C ratio because the lighter ^{12}C isotope is preferentially used for microbial respiration during decomposition (Nadelhoffer and Fry. 1988, Ågren et al. 1996). By contrast, we expected a negative correction factor for the δ^{34}S ratio (i.e. depletion), as this is commonly reported to arise from the activities of sulfate-reducing microbes (Thode 1991). Eq. A.1 was subsequently modified as:

$$\mu_{ij} = \Phi_{\mathrm{SL}ij}\mathrm{SL}_{ij} + \Phi_{\mathrm{LF}ij}\mathrm{LF}_{ij} + \Phi_{\mathrm{PR}ij}\mathrm{PR}_{ij} + \Phi_{\mathrm{PY}ij}\mathrm{PY}_{ij} + \Delta_i \,. \qquad \text{(eq. A.2)}$$

A.1.4 Selection of environmental variables

We characterized sediments with 16 environmental variables in addition to the 4 described in the main text to find those that exhibited the greatest among-site variation. 14 of these variables were water chemistry parameters measured in 24 surface grab samples taken throughout the ice-free season and filtered at 0.45 μm. Total elemental concentrations (magnesium, potassium, sodium, calcium, chromium, cobalt, copper, iron, nickel, zinc, phosphorus), total Kjeldahl nitrogen concentration, sulphate concentration, and conductivity were analyzed by the Ontario Ministry of the Environment and Climate Change using standard methods (OMOE 1983). We also considered the % sulphur content of the sediment samples taken for the isotopic mixing model. Finally, we characterized the particle size distributions of organic material from these same sediment cores. The percentage of material on a dry-mass basis in four size

classes (<0.063, 0.063-1, 1-2, >2 mm) was determined after sieving material as described in Tanentzap et al. (2017).

We selected variables to use in our analyses that displayed statistically significant variation among sites according to a one-way analysis of variance. These variables were OM quantity, % sulfur, pH, and C:N ratio (Table A.1). Additionally, we used a chi-square goodness of fit to compare % terrestrial C among sites. We could not use an ANOVA as the isotopic mixing model estimated a single mean value per site from a posterior distribution. This test showed that % terrestrial C also varied significantly among sites ($X^2 = 248.3$, Df = 7, $p < 0.001$).

We checked for multicollinearity between our five remaining variables with Pearson's correlation tests prior to performing the analyses described in the main text. We found that only % sulfur and OM quantity were highly correlated ($\varrho = 0.87$; Table A.2). Considering that OM quantity varied more among sites (Table A.1), we retained it and discarded % sulfur from our analyses.

A.1.5 DNA extraction, library preparation, and sequencing

DNA concentrations of each extraction were measured prior to preparing amplicon sequencing libraries using the PicoGreen dsDNA quantitation kit (Molecular Probes, Invitrogen, Eugene, OR, USA). Amplicon sequencing libraries were constructed with primers targeting the hyper-variable V3-V4 regions of the 16S rRNA gene (341F (5'-CCTACGGGNGGCWGCAG-3') - 805R (5'-GACTACHVGGGTATCTAATCC -3'), Klindworth et al. 2013). During DNA extraction and PCR amplification, four negative controls (reagent blanks) were used on each 96-well plate. The negative controls revealed no amplification in the targeted region and were therefore discarded prior to sequencing. We then used a dual indexing strategy consisting of adding unique 8 base indices to the primers to allow multiplexing of pooled libraries, along with Illumina

adapter overhang nucleotide sequences to permit binding to the flow cell (Illumina, San Diego, CA, USA). Samples were first amplified with 1 μL of forward and reverse primers (10 nM each), 12.5 μL of 2x KAPA HiFi HotStart ReadyMix (KAPA Biosystems), and 11.5 μL of microbial DNA (5 ng μL^{-1}) in a total volume of 25 μL with the following cycling conditions: initial denaturation of 3 min at 95°C, 25 cycles at 95°C for 30 s, 55°C for 30 s, 72°C for 30 s, and final elongation of 5 min at 72°C. Libraries were purified using 20 μL of Agencourt AMPure XP beads (Beckman Coulter Genomics, Indianapolis, IN). Samples were then amplified again in order to add the Nextera XT Index primers (Illumina). 5 μL of Nextera XT Index Primer 1 (N7XX) and 5 μL of Nextera XT Index Primer 2 (S5XX), 25 μL of 2x KAPA HiFi HotStart ReadyMix, 5 μL of DNA and 10 μL of PCR-grade water in a total volume of 50 μL with the following cycling conditions: initial denaturation of 3 min at 95°C, 8 cycles at 95°C for 30 s, 55°C for 30 s, 72°C for 30 s, and final elongation of 5 min at 72°C. Final libraries were purified using 56 μL of Agencourt AMPure XP beads. Amplicons were quantified on Bioanalyzer DNA 1000 chips (Agilent, Santa Clara, CA, USA) and pooled in equimolar concentrations into a single sample. The final concentration of the library was determined using the PicoGreen dsDNA quantitation kit and was then sequenced on an Illumina MiSeq platform using the MiSeq Reagent Kit v3 (600 cycles, paired-end).

Tagmented libraries for shotgun sequencing were amplified with 5 μL of each Nextera XT Index Primer 1 (N7XX) and Primer 2 (S5XX) and 15 μL of Nextera PCR Master Mix under the following cycling conditions: initial denaturation of 3 min at 72°C followed by 30 s at 95°C, 12 cycles at 95°C for 10 s, 55°C for 30 s, 72°C for 30 s, and final elongation of 5 min at 72°C. Final libraries were purified using 30 μL of Agencourt AMPure XP beads. Libraries were quantified on a Qubit and on Bioanalyzer HS DNA chips (Agilent) and pooled in equimolar concentrations into a single sample.

Sequencing was carried out on an Illumina NextSeq platform using a NextSeq 500/550 Mid Output Kit v2 (300 cycles, paired-end).

A.1.6 Bioinformatics analysis

We sorted the raw reads obtained from amplicon sequencing into groups according to their indices. Indices and primer sequences were then removed, and only sequences with an exact match to the adapter and index primers were kept. The resulting sequences were processed according to the following conditions using mothur version 1.39.5 (Schloss et al. 2009): minimum average quality of 25, no ambiguous bases, length between 400 and 500 bp, and homopolymers no longer than 8 bp. mothur is commonly used for amplicon sequence processing and was made popular by its speed and excellent accompanying documentation (Wooley et al. 2010, Nilakanta et al. 2014, Pollock et al. 2018). Filtered sequences were aligned to the SILVA reference database (Quast et al. 2012) with the kmer searching algorithm and kmer size set to 8 (Schloss et al. 2009). The kmer algorithm is an alignment independent technique that has been shown to outperform other searching algorithms (e.g. blastn and suffix tree searching) both in its ability to find the best template sequence and in its speed (Schloss et al. 2009). Sequences that did not match the reference alignment in the expected positions were discarded. A pre-clustering algorithm was further used to de-noise sequences within each sample and the resulting sequences were screened for chimeras using UCHIME (Edgar 2010). A Bayesian classifier was then used to classify each sequence against the Greengenes taxonomy database (DeSantis et al. 2006), with an 80% pseudobootstrap confidence score. Sequences that did not classify to the level of kingdom or that classified as *Archaea*, *Eukaryota*, chloroplasts, or mitochondria were removed. Finally, the reads were clustered into operational taxonomic units (OTUs) at

97% sequence identity using UCLUST (Edgar 2010). Any read sequenced fewer than six times was removed from subsequent analyses.

Raw sequences from the shotgun sequencing were processed following a modified version of the European Molecular Biology Laboratory-EBI pipeline version 3.0 (Mitchell et al. 2016). The advantage of processing data following the EMBL-EBI pipeline is that they can then be compared more accurately to other datasets processed similarly. The SeqPrep tool (https://github.com/jstjohn/SeqPrep, version 1.1) was used to merge paired-end overlapping reads, Trimmomatic (Bolger et al. 2014; version 0.35) was used to trim low quality ends and sequences with >10% undetermined nucleotides, and <100 nucleotides were removed using Biopython (Cock et al. 2009; version 1.65). Non-coding RNAs were identified and masked with HMMER (http://hmmer.org; version 3.1b1). Sequences were functionally annotated by predicting coding sequences (pCDS) above 60 nucleotides with FragGeneScan (Rho et al. 2010; version 1.20). Read matches were then generated against pCDS using a subset of databases from InterProScan (Jones et al. 2014; version 5.19-58.0) and summarized using the Gene Ontology terms.

Table A.1 | Among-site variation in 20 environmental variables. We fitted ANOVAs to each variable with site as the sole explanatory variable. Variables were ordered by decreasing F value. To achieve normality, sodium, zinc, and C:N ratio were log transformed; total Kjeldhal nitrogen, cobalt, and iron were fourth-root transformed; total phosphorus, copper and nickel were reciprocally transformed; and particle size percentages, % sulfur, % terrestrial C, and OM quantity (% loss on ignition) were logit transformed. For particle size, the ANOVA was nested within four size classes measured on each sample. Df = degrees of freedom.

Response variable	Explanatory variables	Df	Sums of squares	Mean Squares	F value	Pr(>F)
OM quantity	Site	7	119.3	17.04	49.49	<0.001 ***
	Residuals	89	30.65	0.34		
% sulfur	Site	7	121.8	17.40	45.97	<0.001 ***
	Residuals	86	32.55	0.38		
pH	Site	7	6.35	0.91	10.96	<0.001 ***
	Residuals	89	7.36	0.08		
C:N ratio	Site	7	2.33	0.33	3.38	0.003 **
	Residuals	86	8.49	0.1		
Total Kjeldahl Nitrogen	Site	7	0.01	2×10^{-3}	1.62	0.201
	Residuals	16	0.02	1×10^{-3}		
Calcium	Site	7	0.21	0.03	1.61	0.203
	Residuals	16	0.30	0.02		
Particle size	Site	6	11.36	1.89	1.24	0.308
	Residuals	38	57.98	1.53		
	Error	3	117.20	39.08		
Copper	Site	7	1.08	0.15	1.23	0.343
	Residuals	16	2.00	0.13		

Nickel	Site	7	0.52	0.08	1.21	0.355
	Residuals	16	0.99	0.06		
Magnesium	Site	7	0.04	0.01	1.14	0.390
	Residuals	16	0.07	4 x 10^{-3}		
Iron	Site	7	0.25	0.04	0.97	0.485
	Residuals	16	0.58	0.04		
Cobalt	Site	7	0.03	4 x 10^{-3}	0.97	0.487
	Residuals	16	0.06	4 x 10^{-3}		
Zinc	Site	7	0.16	0.02	0.97	0.488
	Residuals	16	0.37	0.02		
Total Phosphorus	Site	7	67184	9598	0.92	0.516
	Residuals	16	166666	10417		
Chromium	Site	7	9.29 x 10^{-8}	1.33 x 10^{-8}	0.64	0.719
	Residuals	16	3.33 x 10^{-7}	2.08 x 10^{-8}		
Sodium	Site	7	0.02	2.00 x 10^{-3}	0.40	0.886
	Residuals	16	0.09	0.01		
Conductivity	Site	7	6.62	0.95	0.30	0.943
	Residuals	16	50.3	3.14		
Sulphate	Site	7	0.33	0.05	0.26	0.960
	Residuals	16	2.86	0.18		
Potassium	Site	7	1.00 x 10^{-3}	1.30 x 10^{-4}	0.10	0.997
	Residuals	16	0.02	1 x 10^{-3}		

Significance levels: * = <0.05, ** = <0.01, *** = <0.001

Table A.2 | Pearson's correlation matrix for variables displaying among-site variation. For each variable, we correlated site-level means to ensure data of matching length.

	pH	OM quantity	% terrestrial C	C:N ratio	% sulfur
pH	1				
OM quantity	0.22	1			
% terrestrial C	-0.57	0.10	1		
C:N ratio	0.26	0.44	-0.35	1	
% sulfur	-0.10	0.87	0.37	0.35	1

Table A.3 | Environmental gradients across our study lake. We measured pH and OM quantity (% loss on ignition) for each sediment core used to measure ecosystem functioning (total n = 97) and % terrestrial carbon (C) and C:N ratio for each sediment sample taken for the isotopic mixing model (total n = 94).

Variable	Minimum	Median	Maximum
pH	4.71	5.95	6.67
OM quantity	0.61	13.19	73.83
% terrestrial C	1.10	6.60	81.80
C:N ratio	12.02	16.41	17.35

Table A.4 | Gene Ontology (GO) categories associated with each of our four subsets of functional genes. Each subset is involved in different aspects of terrestrial OM decomposition.

Functional gene subset	GO categories
Hydrolase	"cellulase activity"
	"glucosidase activity"
	"xylan 1,4-beta-xylosidase activity"
Oxidase	"peptidase activity"
	"phosphatase activity"
	"peroxidase activity"
	"nitrate reductase activity"
	"catechol oxidase activity"
	"catechol 1,2-dioxygenase activity"
Carbohydrate metabolism	"glycolytic process"
	"polysaccharide catabolic process"
Aromatic compound catabolism	"aromatic compound catabolic process"

Table A.5 | Model selection statistics and parameter estimates for models of CO_2 production with each of four functional gene subsets. AICc weight (*w*) is the proportion of support for a given model out of the candidate set and marginal R^2 is a measure of variance explained by model fixed effects only. Mean estimates and standard errors (SE) are given for the fixed effect predictors of each model. Variance is given for the site-level random effect. Bolded values have 95% CI that exclude 0 and are considered statistically significant. All abundance measures are normalized abundances.

Predictors		Functional gene abundance			
		Hydrolases	***Oxidases***	***Carbohydrate metabolism***	***Aromatic compound catabolism***
.	AICc	102.70	103.40	102.40	103.90
	w	0.29	0.21	0.34	0.16
	R^2	0.58	0.68	0.62	0.60
Intercept	Mean	4.70	4.94	4.75	4.73
	SE	0.12	0.17	0.13	0.15
pH	Mean	-0.16	**-0.43**	-0.15	-0.20
	SE	0.12	**0.16**	0.12	0.14
OM quantity	Mean	**0.42**	**0.62**	**0.60**	**0.48**
	SE	**0.20**	**0.18**	**0.24**	**0.21**
% terrestrial C	Mean	-0.16	-0.04	-0.08	-0.10
	SE	0.18	0.21	0.25	0.26
C:N ratio	Mean	-0.04	-0.06	-0.05	0.07
	SE	0.18	0.19	0.20	0.20

Bacterial abundance	Mean	-0.32	-0.11	-0.05	-0.12
	SE	0.28	0.22	0.21	0.27
Taxonomic diversity	Mean	**0.47**	**0.68**	**0.45**	**0.44**
	SE	**0.23**	**0.24**	**0.20**	**0.20**
Functional diversity	Mean	0.00	-0.27	-0.16	-0.25
	SE	0.24	0.18	0.18	0.18
Functional gene abundance	Mean	-0.21	**0.51**	-0.34	0.20
	SE	0.25	**0.24**	0.24	0.19
Functional gene abundance × C:N	Mean	0.16	0.30	0.23	-0.17
	SE	0.17	0.17	0.18	0.20
Functional gene abundance × % terrestrial C	Mean	0.05	**0.59**	0.02	0.02
	SE	0.17	**0.27**	0.14	0.15
Site	Variance	0.03	0.09	0.04	0.05

A.2 Supplementary results

Table A.6 | Sequencing and taxonomic output of both the shotgun (n = 22) and the amplicon (n = 97) sequencing datasets. Percentage abundance of different phyla is based on DESeq transformed abundances.

	Library	
	Shotgun	**Amplicon**
# of sequences	923,560	3,831,950
# of OTUs	22,899	27,027
# of families	531	551
# of phyla	62	64
Top 3 phyla (% abundance)		
Proteobacteria	22	22
Chloroflexi	11	11
Acidobacteria	9	9

Table A.7 | Community structure across our study lake in the 22 samples used in the path analysis. We measured taxonomic and functional diversity as Shannon's H' and the remaining metrics as normalized abundances (DESeq transformed). SE = standard error.

Diversity metric	Minimum	Mean (± SE)	Maximum
Bacterial abundance	1901.93	2046.91 (± 18.51)	2228.81
Taxonomic diversity	3.83	4.42 (± 0.04)	4.76
Functional diversity	5.03	5.05 (± 0.003)	5.08
Hydrolases	25.58	26.13 (± 0.07)	26.78
Oxidases	55.68	56.14 (± 0.05)	56.65
Carbohydrate metabolism	22.17	22.57 (± 0.05)	23.10
Aromatic compound catabolism	9.06	9.92 (± 0.07)	10.31

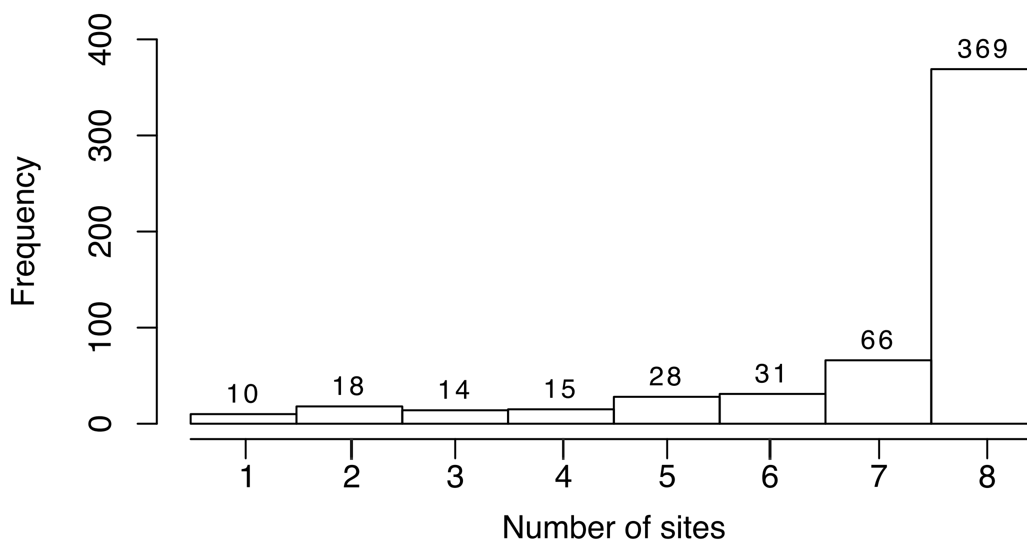

Fig. A.2 | Frequency distribution of family-level presence across sites. Numbers above each bar indicate the number of unique taxa found between 1 to 8 sites.

Table A.8 | Partitioning the effect of the past and present environment on bacterial community composition at the family level. (a) Sum of the eigenvalues of the five canonical correspondence analysis (CCA) axes (pH, OM quantity, % terrestrial C, C:N ratio, and site; i.e. "Inertia") and their contribution to the total variance (i.e. "Proportion") in community composition. (b) Variation in bacterial community composition explained by PERMANOVA (999 permutations). Df = degrees of freedom. (c) Sum of the eigenvalues of the four redundancy analysis (RDA) axes (pH, OM quantity, % terrestrial C, and C:N ratio; i.e. "Inertia") conditional on site and their contribution to the total variance (i.e. "Proportion") in community composition.

a) CCA – Partitioning of correlations

Explained variance	Inertia	Proportion
Total	0.06	1.00
Constrained	0.02	0.40
Unconstrained	0.03	0.60

b) PERMANOVA output table

	Df	Sums of Squares	F	R^2	Pr(>F)
pH	1	0.01	1.84	0.02	0.06
OM quantity	1	0.01	2.14	0.02	0.06
% terrestrial C	1	0.03	5.64	0.05	0.001 ***
C:N ratio	1	0.03	4.61	0.04	0.001 ***
Site	1	0.02	3.92	0.03	0.01 **

Residuals	91	0.55	0.77
Total	96	0.71	1.00

Significance levels: * = <0.05

c) Partial RDA – Partitioning of correlations

Explained variance	Inertia	Proportion
Total	493	1.00
Conditional	37	0.07
Constrained	74	0.15
Unconstrained	382	0.78

Table A.9 | Model averaged parameter effects used to predict CO_2 production. 95% confidence intervals (CI) were averaged across competing models for each predictor, while values for functional gene abundance, its interactions with the environment and the model intercept were specific to the oxidase functional gene subset, which was the only gene subset with a statistically significant effect (Table A.5). Bolded predictors were those with 95% CI that did not overlap 0 and were considered statistically significant. All abundance measures are normalized abundances.

Fixed predictors	**2.5%**	**97.5%**
Oxidase-derived intercept	4.60	5.25
pH	**-0.36**	**-0.07**
OM quantity	**0.30**	**0.80**
% terrestrial C	-0.33	0.15
C:N ratio	-0.22	0.18
Bacterial abundance	-0.45	0.12
Taxonomic diversity	**0.28**	**0.74**
Functional diversity	-0.37	0.06
Oxidase gene abundance	**0.02**	**0.98**
Oxidase gene abundance × C:N ratio	-0.03	0.63
Oxidase gene abundance × % terrestrial C	**0.03**	**1.14**

Table A.10 | Mean parameter estimates and 95% confidence interval (CI) for predictors in models with each of the four measures of community structure as a response. Bolded predictors were those with 95% CI that did not overlap 0 and were considered statistically significant. SE = standard error. All abundance measures are normalized abundances.

Predictor		**Response variable**			
		Bacterial abundance	***Taxonomic diversity***	***Functional diversity***	***Oxidase abundance***
Intercept	Mean	7.62	4.43	5.05	56.14
	SE	0.01	0.04	<0.01	0.04
	95% CI	(7.60, 7.64)	(4.35, 4.51)	(5.05, 5.06)	(56.05, 56.22)
pH	Mean	0.00	-0.02	0.00	0.06
	SE	0.01	0.03	<0.01	0.04
	95% CI	(-0.02, 0.01)	(-0.09, 0.05)	(<-0.01, 0.01)	(-0.02, 0.14)
OM quantity	Mean	-0.02	0.02	0.00	-0.05
	SE	0.01	0.05	<0.01	0.05
	95% CI	(-0.04, 0.00)	(-0.08, 0.12)	(<-0.01, 0.01)	(-0.16, 0.07)
% terrestrial C	Mean	-0.02	-0.01	0.00	-0.03
	SE	0.01	0.05	<0.01	0.06
	95% CI	(-0.04, 0.00)	(-0.13, 0.10)	(-0.01, <0.01)	(-0.14, 0.08)
C:N ratio	Mean	-0.01	0.00	0.00	-0.03
	SE	0.01	0.05	<0.01	0.06
	95% CI	(-0.04, 0.02)	(-0.12, 0.10)	(<-0.01, 0.01)	(-0.14, 0.10)
Bacterial abundance	Mean		**0.17**	0.01	0.12
	SE		**0.05**	<0.01	0.07
	95% CI		**(0.07, 0.26)**	(<-0.01, 0.01)	(-0.03, 0.27)

Taxonomic diversity	Mean			**0.01**	**-0.19**
	SE			**<0.01**	**0.06**
	95% CI			**(<0.01, 0.01)**	**(-0.30, -0.07)**
Functional diversity	Mean				0.10
	SE				0.06
	95% CI				(-0.01, 0.22)
Standard deviation in site	Mean	0.03	0.08	<0.01	0.07
	95% CI	(0.01, 0.04)	(0.03-0.16)	(<0.001-0.01)	(0.03-0.18)

Table A.11 | No significant effect of CO_2 production in predicting the four measures of community structure. We checked that there were no missing linkages in our model by including CO_2 production as a predictor of the four measures of community structure. This missing linkage had no statistically significant effect in models predicting measures of community structure (see Table A.10), suggesting that there was no direct influence of ecosystem function on microbial diversity and abundance. All abundance measures are normalized abundances.

Predictor		Response variable			
		Bacterial abundance	***Taxonomic diversity***	***Functional diversity***	***Oxidase abundance***
Intercept	Mean	7.62	4.43	5.05	56.14
	SE	0.01	0.04	<0.01	0.04
	95% CI	(7.60, 7.64)	(4.35, 4.50)	(5.05, 5.06)	(56.06, 56.22)
CO_2 production	Mean	0.00	0.04	0.00	0.03
	SE	0.01	0.04	<0.01	0.05
	95% CI	(-0.01, 0.02)	(-0.04, 0.11)	(-0.01, <0.01)	(-0.05, 0.12)
pH	Mean	0.00	-0.01	0.00	0.07
	SE	0.01	0.04	<0.01	0.04
	95% CI	(-0.02, 0.01)	(-0.08, 0.06)	(<-0.01, <0.01)	(-0.02, 0.16)
OM quantity	Mean	-0.02	-0.01	0.00	-0.08
	SE	0.01	0.04	<0.01	0.06
	95% CI	(-0.04, 0.00)	(-0.12, 0.10)	(<-0.01, 0.01)	(-0.21, 0.04)
% terrestrial C	Mean	-0.02	-0.01	0.00	-0.01
	SE	0.01	0.05	<0.01	0.06
	95% CI	(-0.04, 0.01)	(-0.11, 0.10)	(<-0.01, <0.01)	(-0.13, 0.11)

C:N ratio	Mean	-0.01	0.00	0.00	-0.03
	SE	0.01	0.05	<0.01	0.06
	95% CI	(-0.03, 0.02)	(-0.10, 0.11)	(<-0.01, 0.01)	(-0.14, 0.10)
Bacterial abundance	Mean		**0.16**	0.00	0.12
	SE		**0.04**	<0.01	0.07
	95% CI		**(0.07, 0.26)**	(<-0.01, 0.01)	(-0.02, 0.28)
Taxonomic diversity	Mean			**0.01**	**-0.21**
	SE			**<0.01**	**0.06**
	95% CI			**(<0.01, 0.01)**	**(-0.32, -0.08)**
Functional diversity	Mean				0.11
	SE				0.06
	95% CI				(-0.01, 0.23)
Standard deviation in site	Mean	0.03	0.07	<0.01	0.07
	95% CI	(0.01, 0.04)	(0.03-0.15)	(<0.01-0.01)	(0.03-0.17)

Appendix B

Diversity-stability relationships in lake sediment microbiomes

B.1 Supplementary methods

B.1.1 DNA extraction, library preparation, and sequencing

Microbial communities were sampled from surface sediments (~ top 5 cm) and were immediately placed into individual sterile sample bags. Sediments were then freeze-dried and stored at -20˚C to stabilize the microbial communities (Miller et al. 1999).

To obtain taxonomic information, we carried out next-generation sequencing. DNA was extracted from each mesocosm in duplicate using a Power Soil DNA Isolation Kit (MoBio Laboratories Inc., Carlsbad, CA, USA) according to the manufacturer's instructions. Duplicates were pooled for downstream analysis. The DNA concentration of each sample was measured using the PicoGreen dsDNA quantitation kit (Molecular Probes, Invitrogen, Eugene, OR, USA). Sequencing libraries were prepared with 1 ng of genomic DNA per sample using the Nextera XT DNA Sample Preparation and dual-barcoding with Nextera XT Indexes (Illumina, San Diego, CA, USA) following the manufacturer's instructions. Libraries were quantified on a Qubit 3.0 Fluorometer (ThermoFisher Scientific, Waltham, MA, USA) and on a Bioanalyzer HS DNA chip (Agilent, Santa Clara, CA, USA) and pooled in equimolar concentrations into a single sample. Samples were sequenced on a NextSeq (Illumina) using a NextSeq 500/550 Mid Output Kit v2 (300 cycles, paired-end).

B.1.2 Bioinformatics analysis

Raw sequences were processed at a read depth of approximately 3.3 million reads per sample following a modified version of the European Molecular Biology Laboratory-EBI pipeline version 3.0 (Mitchell et al. 2016). The SeqPrep tool version 1.1 (https://github.com/jstjohn/SeqPrep) was used to merge paired-end overlapping reads and Trimmomatic version 0.35 (Bolger et al. 2014) was used to trim low quality ends. Sequences with >10% undetermined nucleotides and <100 nucleotides were removed using Biopython version 1.65 (Cock et al. 2009). ncRNAs were identified and removed with HMMER version 3.1b1 (http://hmmer.org). Sequences were taxonomically annotated using QIIME version 1.9.1 (Caporaso et al. 2010). Representative 16S sequences were classified with the SILVA reference database (release 128) at 97% sequence identity following the open-reference OTU picking method with reverse strand matching enabled. We removed 14 OTUs out of 7,696 that were found in the negative control with a relative abundance >1%, 73% of which belonged to the Euryarchaeota (as per Flores et al. 2012).

B.2 Supplementary results

Table B.1 | Model outputs of linear regressions testing ultimate and proximal drivers of CO_2 and community stability. EF = ecosystem function; n.s. = non-significant (p>0.05); SE = standard error.

Measure of EF	Response	Predictor	Slope (SE)	p-value	R^2
CO_2 production	Stability	OTU richness	0.16 (0.07)	0.041	0.17
		OM quantity	-0.08 (0.09)	n.s.	
		OM quality	0.07 (0.08)	n.s.	
		OTU richness × OM quantity	0.09 (0.10)	n.s.	
		OTU richness × OM quality	-0.08 (0.08)	n.s.	
		Asynchrony	-0.03 (0.09)	n.s.	
		Bay	0.06 (0.13)	n.s.	
	σ	OTU richness	-0.15 (0.06)	0.016	0.30
		OM quantity	0.19 (0.05)	0.002	
		OM quality	0.02 (0.06)	n.s.	
		OTU richness × OM quantity	-0.18 (0.07)	0.024	
		OTU richness × OM quality	0.07 (0.06)	n.s.	
		Bay	0.03 (0.10)	n.s.	
	μ	OTU richness	0.02 (0.06)	n.s.	0.28
		OM quantity	0.10 (0.06)	n.s.	
		OM quality	0.09 (0.07)	n.s.	
		OTU richness × OM quantity	-0.08 (0.08)	n.s.	
		OTU richness × OM quality	0.00 (0.06)	n.s.	
		Bay	0.09 (0.10)	n.s.	

Microbial abundance	Stability	OTU richness	0.10 (0.14)	n.s.	0.64
		OM quantity	0.28 (0.16)	n.s.	
		OM quality	-0.14 (0.15)	n.s.	
		OTU richness × OM quantity	0.24 (0.18)	n.s.	
		OTU richness × OM quality	0.00 (0.14)	n.s.	
		Asynchrony	0.63 (0.16)	<0.001	
		Bay	0.36 (0.24)	n.s.	
	σ	Asynchrony	-0.64 (0.15)	<0.001	0.39
		Bay	-0.17 (0.29)	n.s.	
	μ	Asynchrony	0.09 (0.08)	n.s.	<0.01
		Bay	0.10 (0.16)	n.s.	

Table B.2 | Model outputs of linear regressions testing the effects of diversity and the environment on asynchrony. n.s. = non-significant (p>0.05)); SE = standard error.

Response	Predictor	Slope (SE)	p-value	R^2
Asynchrony	OTU richness	-0.04 (0.06)	n.s.	0.37
	OM quantity	0.17 (0.05)	0.006	
	OM quality	0.02 (0.06)	n.s.	
	OTU richness × OM quantity	-0.00 (0.07)	n.s.	
	OTU richness × OM quality	-0.07 (0.06)	n.s.	
	Bay	0.02 (0.10)	n.s.	

Appendix C

Changes in lake sediment microbiomes in their early development

C.1 Supplementary methods

C.1.1 DNA extraction, library preparation, and sequencing

Surface sediment grabs were immediately placed into individual sterile sample bags and then freeze-dried at -20˚C to stabilize the communities (Miller et al. 1999).

To obtain taxonomic information, we first extracted DNA in duplicate from each sample using the Power Soil DNA Isolation Kit (MoBio Laboratories Inc., Carlsbad, CA, USA) according to the manufacturer's instructions. We also extracted DNA from the original t-OM material each in triplicate. Duplicate extractions from sediment grabs were pooled for downstream analysis whilst triplicate extractions from the original t-OM material were processed individually. The DNA concentration of each sample was measured using the PicoGreen dsDNA quantitation kit (Molecular Probes, Invitrogen, Eugene, OR, USA) prior to downstream library preparation.

Sequencing libraries were constructed using the resulting 249 DNA samples and three negative controls comprised of double-distilled water. Amplicon sequencing libraries were constructed using a two-step PCR process. Firstly, bacterial and archaeal DNA was amplified with primers targeting the hyper-variable V3-V4 regions of the universal 16S rRNA gene: 341F (5'- CCTAYGGGRBGCASCAG-3') - 806R (5'-GGACTACNNGGGTATCTAAT-3') (Yu et al. 2005), and the ITS1 region was used to

target fungi with the primers ITS1f (5'-CTTGGTCATTTAGAGGAAGTAA-3', Gardes & Bruns 1993) and ITS2 (5'-GCTGCGTTCTTCATCGATGC-3',White et al. 1990), each also including an overhang sequence. A second PCR used primers with the overhang tails, a 6 nucleotide (nt) index sequence and Illumina capture sequences to permit binding to the flow cell as in Campbell et al. (2015), and were synthesized using the Trugrade process (IDT, Leuven, Belgium). This dual indexing strategy was used to allow multiplexing of pooled libraries.

For the PCR, samples were first amplified with 2 μL of forward and 2 μL of reverse primers (1 μM each), 10 μL of Qiagen Multiplex PCR Master Mix (Qiagen, UK), 4 μL of sterile double-distilled water and 2 μL of microbial DNA (10 ng/μL) in a total volume of 20 μL with the following cycling conditions: initial denaturation of 15 min at 95°C, 35 cycles at 94˚C for 30 s, 55˚C for 45 s, 72˚C for 30 s, and final elongation of 10 min at 72˚C. Samples were amplified again to add the Trugrade index primers with 2 μL of Forward Index Primer 1 (i5, 1 μM) and 2 μL of Reverse Index Primer 2 (i7, 1 μM), 10 μL of Qiagen Multiplex PCR Master Mix and 8 μL of template DNA in a total volume of 22 μL with the following cycling conditions: initial denaturation of 15 min at 95°C, 10 cycles at 98˚C for 10 s, 65˚C for 30 s, 72˚C for 30 s, and final elongation of 5 min at 72˚C. Samples were quantified on a FLUOstar OPTIMA plate reader at 545 nm (BMG Labtech, Aylesbury, UK) and pooled in groups of 8 in equimolar quantities (150 ng). Final libraries (50 μL) were purified using a first round of 25 μL of Agencourt AMPure XP beads (Beckman Coulter Genomics, Indianapolis, IN, USA), and a second round of 67.5 μL of beads. Amplicons were quantified on a Quantstudio 12k Flex Real-Time PCR system (Applied Biosystems, Warrington, UK) with 6 μL of KAPA SYBR FAST mix and primers (KAPA Biosystems, Wilmongton, MA, USA) and 2 μL of PCR-grade water with the following cycling conditions: 95˚C for 5 min, 35 cycles at 95˚C for 30s and 60˚C for 45 sec. The

size of the amplicons was checked on an Agilent 4200 TapeStation (Agilent, Santa Clara, CA, USA) and pooled in equimolar concentrations into a single sample. The final concentration of the library was determined using PicoGreen dsDNA quantitation and was then sequenced on an Illumina MiSeq sequencer (600 cycles, paired-end) using the MiSeq Reagent Kit v3 (Illumina).

C.1.2 Bioinformatics analysis

We used DADA2 to infer the amplicon sequence variants (ASVs) present in each sample (Callahan et al. 2016a). First, we removed the 16S and ITS primer adapters from our sequences using *cutadapt* (Martin 2011). We then followed a slightly modified version of the DADA2 Workflow for Big Data (Callahan et al. 2016b). As 16S and ITS samples were sequenced on separate runs, their quality scores and error rates were estimated separately to account for run-to-run variability. For samples targeting the 16S primer, we trimmed forward and reverse reads at 280 nt and 200 nt respectively, and filtered them such that there were no ambiguous bases and that each read had a maximum of 8 expected errors, as determined by checking their quality scores. We did not trim samples targeting the ITS primers due to its variable length (Nilsson et al. 2008). Forward and reverse reads were filtered such that they had a maximum of 2 and 4 expected errors, as determined by their respective quality scores, and that they had no ambiguous bases. Using the run-specific error rates, we thus inferred ASVs for the forward and reverse reads of both the 16S and ITS runs, after which we merged the read pairs and removed ASVs identified as chimeras. Finally, taxonomy was assigned to the assembled ASVs using the RDP naïve Bayesian classifier implemented in DADA2 against the SILVA reference database version 128 for the 16S reads (Quast et al. 2013) and the UNITE reference database version 7.2 for the ITS reads (Kõljalg et al. 2013).

C.2 Supplementary results

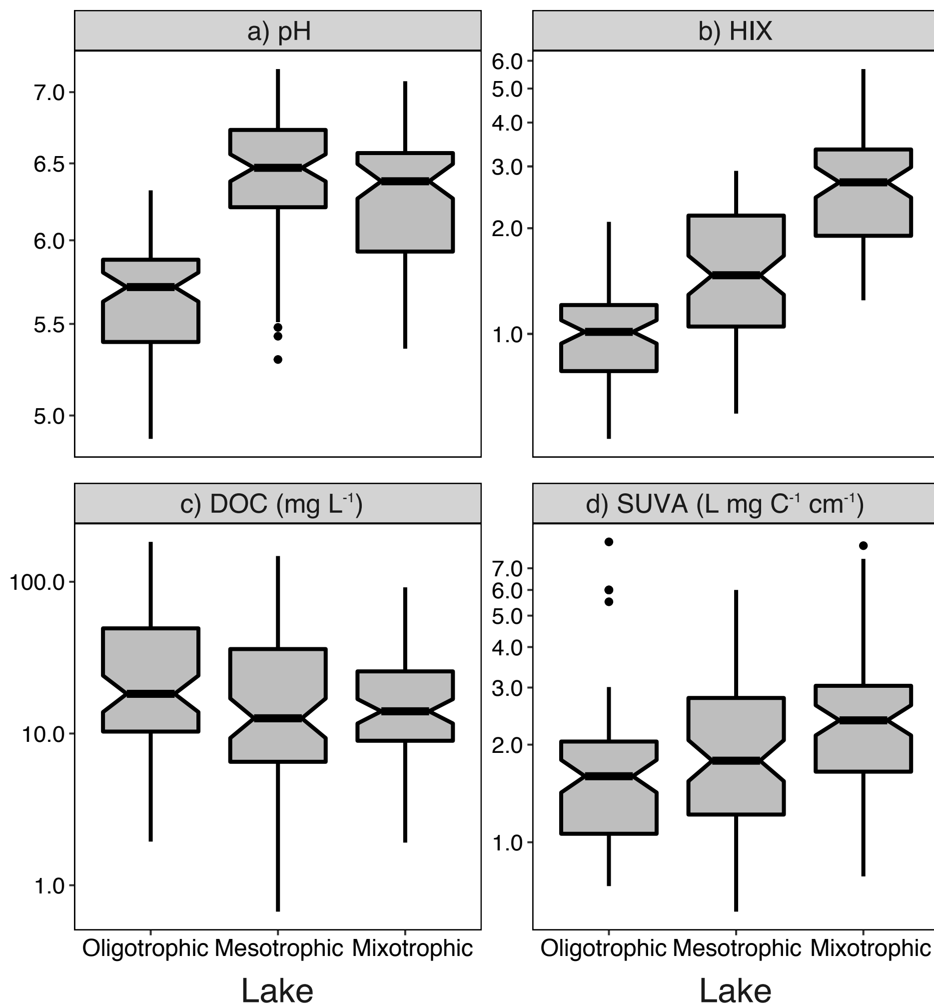

Fig. C.1 | Sediment pore water conditions differed between mesocosms in the three lakes. a) The oligotrophic lake had lower pH than the other two lakes, b) HIX increased in darker, nutrient-rich lakes, c) DOC was lower in the mixotrophic lake compared to the oligotrophic lake, and d) SUVA was higher in the mixotrophic lake. Non-overlapping notches indicate differences in the two medians based on 95% confidence intervals (Chambers et al. 1983). The upper and lower whiskers extended 1.5 times the interquartile range, with points outside of this range plotted.

Table C.1 | Model output of linear regression testing rates of compositional change from the original leaf material per day. Mixotrophic status is in the intercept. Day refers to the sampling day of experiment.

Predictor	Mean effect (standard error)	*t*-value	*P*-value
Intercept	5.99×10^{-2} (4.64×10^{-3})	12.91	<0.001 ***
Sampling day	4.35×10^{-4} (1.21×10^{-4})	3.58	<0.001 ***
Oligotrophic	1.63×10^{-4} (6.13×10^{-3})	0.03	0.979
Mesotrophic	7.72×10^{-4} (6.73×10^{-3})	0.11	0.910
Day:Oligotrophic	-6.83×10^{-5} (1.62×10^{-4})	-0.42	0.674
Day:Mesotrophic	-2.72×10^{-4} (1.82×10^{-4})	-1.50	0. 136

Significance levels: * = <0.05, ** = <0.01, *** = <0.001

Table C.2 | Environmental fit (envfit) results of the associations between environmental conditions and compositional shifts in the NMDS ordination. Significance was determined with 999 permutations.

Predictor	Axis 1	Axis 2	R^2	*P*-value
HIX	-0.98	-0.17	0.04	0.009 **
SUVA	-0.86	0.51	0.01	0.290
pH	-0.99	-0.17	0.05	0.001 ***
DOC	0.80	-0.60	0.01	0.385

Significance levels: ** = <0.01, *** = <0.001

Table C.3 | ANOVAs comparing full model to reduced models without the three main effects and their interactions with date. Each row represents the respective removal of the three environmental predictors (main effects and interaction terms) from the full model.

Full model:

Dissimilarity ~ Sampling Day x Lake Comparison + Sampling Day x t-OM quality + Sampling Day x t-OM quantity + Bay

	Df	**Sums of Squares**	***F* value**	***P*-value**	**R^2**
Full Model	228				0.42
- Lake Comparison	232	-3.79×10^{-3}	8.07	<0.001 ***	0.35
- t-OM Quality	232	-1.81×10^{-3}	3.87	0.005 **	0.39
- t-OM Quantity	232	-1.17×10^{-2}	24.83	<0.001 ***	0.18

Significance levels: ** = <0.01, *** = <0.001

Table C.4 | Model output of linear regression predicting changes in community dissimilarity. The oligo-mixo lake comparison, 5% t-OM quantity treatment, and coniferous t-OM quality treatment are in the intercept. Day refers to the sampling day of experiment.

Predictor	Mean effect (standard error)	*t*-value	*P*-value
Intercept	5.52×10^{-2} (2.70×10^{-3})	20.49	<0.001 ***
Sampling day	1.75×10^{-4} (6.74×10^{-5})	2.59	0.010 *
Oligo-Meso	1.08×10^{-2} (2.29×10^{-3})	4.71	<0.001 ***
Meso-Mixo	-4.94×10^{-4} (2.47×10^{-3})	-0.20	0.841
Mixed	3.49×10^{-3} (2.51×10^{-3})	1.39	0.167
Deciduous	-1.04×10^{-3} (2.27×10^{-3})	-0.46	0.646
25%	7.46×10^{-3} (2.56×10^{-3})	2.92	0.004 **
50%	1.10×10^{-2} (2.62×10^{-3})	4.20	<0.001 ***
Bay	-5.79×10^{-4} (1.44×10^{-3})	-0.40	0.688
Day:Oligo-Meso	-8.01×10^{-5} (6.52×10^{-5})	-1.23	0.220
Day:Meso-Mixo	6.63×10^{-5} (6.66×10^{-5})	1.00	0.321
Day:Mixed	-1.25×10^{-4} (6.71×10^{-5})	-1.86	0.064
Day:Deciduous	-1.39×10^{-4} (6.18×10^{-5})	-2.25	0.025 *
Day:25%	1.26×10^{-4} (6.87×10^{-5})	1.84	0.067
Day:50%	1.85×10^{-4} (6.93×10^{-5})	2.66	0.008 **

Significance levels: * = <0.05, ** = <0.01, *** = <0.001

Table C.5 | Archaeal and fungal abundances increased with time while bacterial ones decreased. We fit linear models to assess how each kingdom's abundances (DESeq-transformed read counts) changed with time.

Microbial kingdom	Predictor	Mean effect (standard error)	*t*-value	*P*-value
Archaea	Intercept	3.26 (0.03)	110.12	<0.001 ***
	Sampling day	0.02 (8.01×10^{-4})	24.77	<0.001 ***
Fungi	Intercept	3.42 (0.01)	577.29	<0.001 ***
	Sampling day	1.60×10^{-3} (1.60×10^{-4})	9.95	<0.001 ***
Bacteria	Intercept	3.60 (0.01)	640.51	<0.001 ***
	Sampling day	-4.97×10^{-4} (1.52×10^{-4})	-3.27	<0.01 **

Significance levels: ** = <0.01, *** = <0.001

Table C.6 | Model outputs of linear regressions testing the effects of t-OM quantity, t-OM quality, and lake, respectively, on changes in degrees with time. Separate models were run for each environmental predictor of changes in degrees with time.

Environmental predictor	Treatment predictor	Mean effect (standard error)	*t*-value	*P*-value
t-OM quantity	Intercept	-3.98 (5.51)	-0.72	0.471
	25%	10.62 (2.16)	4.92	<0.001 ***
	50%	11.22 (2.15)	5.22	<0.001 ***
	Bacteria	0.60 (5.44)	0.11	0.850
	Fungi	-5.18 (5.52)	-0.94	0.119
t-OM quality	Intercept	-3.88 (5.07)	-0.76	0.455
	Mixed	5.24 (2.08)	2.52	0.012 *
	Deciduous	0.72 (2.07)	0.35	0.730
	Bacteria	5.44 (5.17)	1.05	0.293
	Fungi	-0.26 (5.24)	-0.05	0.960
Lake	Intercept	-1.31 (6.03)	-0.22	0.829
	Oligotrophic	3.80 (2.21)	1.72	0.086
	Mesotrophic	2.14 (2.41)	0.89	0.375
	Bacteria	4.57 (5.84)	0.78	0.434
	Fungi	-1.96 (5.93)	-0.33	0.741

Significance levels: * = <0.05, *** = <0.001

Table C.7 | Taxonomic identification of the 10 most abundant ASVs.

Code	Kingdom	Phylum	Class	Order	Family	Genus	Species
ASV 1	Archaea	Euryarchaeota	Methanobacteria	Methanobacteriales	Methanobacteriaceae	*Methanobacterium*	-
ASV 2	Archaea	Euryarchaeota	Methanobacteria	Methanobacteriales	Methanobacteriaceae	*Methanobacterium*	-
ASV 3	Bacteria	Firmicutes	Clostridia	Clostridiales	Clostridiaceae	*Clostridium*	-
ASV 4	Fungi	Ascomycota	Sordariomycetes	Hypocreales	Hypocreaceae	*Trichoderma*	-
ASV 5	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Trimorphomycetaceae	*Saitozyma*	*podzolica*
ASV 6	Fungi	Mucoromycota	Umbelopsidomycetes	Umbelopsidales	Umbelopsidaceae	*Umbelopsis*	*isabellina*
ASV 7	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Debaryomycetaceae	*Meyerozyma*	*guilliermondii*
ASV 8	Fungi	Mucoromycota	Umbelopsidomycetes	Umbelopsidales	Umbelopsidaceae	*Umbelopsis*	*isabellina*
ASV 9	Fungi	Ascomycota	Sordariomycetes	Hypocreales	Hypocreaceae	*Trichoderma*	*deliquescens*
ASV 10	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Dipodascaceae	*Dipodascus*	*geotrichum*

References

Åberg J, Wallin MB. Evaluating a fast headspace method for measuring DIC and subsequent calculation of pCO2 in freshwater systems. *Inl Waters* 2014; **4**: 157–166.

Adams RI, Miletto M, Lindow SE, Taylor JW, Bruns TD. Airborne bacterial communities in residences: Similarities and differences with fungi. *PLoS One* 2014; **9**: e91283.

Ågren GI, Bosatta E, Balesdent J. Isotope discrimination during decomposition of organic matter: A theoretical analysis. *Soil Sci Soc Am J* 1996; **60**:1121–1126.

Amaral V, Graeber D, Calliari D, Alonso C. Strong linkages between DOM optical properties and main clades of aquatic bacteria. *Limnol Oceanogr* 2016; **61**: 906–918.

Amir A, McDonald D, Navas-Molina JA, Kopylova E, Morton JT, Zech Xu Z, Kightley EP, Thompson LR, Hyde ER, Gonzalez A, Knight R. Deblur Rapidly Resolves Single-Nucleotide Community Sequence Patterns. *mSystems* 2017; **2**: e00191-16.

Amon RMW, Benner R. Bacterial utilization of different size classes of dissolved organic matter. *Limnol Oceanogr* 1996; **41**: 41–51.

Andam CP, Doroghazi JR, Campbell AN, Kelly PJ, Choudoir MJ, Buckley DH. A latitudinal diversity gradient in terrestrial bacteria of the genus Streptomyces. *MBio* 2016; **7:** e02200-15.

Antwis RE, Griffiths SM, Harrison XA, Aranega-Bou P, Arce A, Bettridge AS, Brailsford FL, deMenezes A, Devaynes A, Forbes KM, Fry EL, Goodhead I, Haskell E, Heys C, James C, Johnston SR, Lewis GR, Lewis Z, Macey MC, McCarthy A, McDonald JE, Florez NLM, O'Brien D, Orland C, Pautasso M, Reid WDK, Robinson HA, Wilson K, Sutherland WJ. Fifty important research questions in microbial ecology. *FEMS Microbiol Ecol* 2017; **93**: 1–10.

Azam F, Fenchel T, Field J, Gray J, Meyer-Reil L, Thingstad F. The Ecological Role of Water-Column Microbes in the Sea. *Mar Ecol Prog Ser* 1983; **10**: 257–263.

Baas-Becking LGM. Geobiologie of inleiding tot de milieukunde. 1934; WP Van Stockum & Zoon: The Hague, Netherlands.

Bahram M, Hildebrand F, Forslund SK, Anderson JL, Soudzilovskaia NA, Bodegom PM, Bengtsson-Palme J, Anslan S, Coelho LP, Harend H, Huerta-Cepas J, Medema MH, Maltz MR, Mundra S, Olsson PA, Pent M, Põlme S, Sunagawa S, Ryberg M, Tedersoo L, Bork P. Structure and function of the global topsoil microbiome. *Nature* 2018; **560**: 233–237.

Baldy V, Chauvet E, Charcosset JY, Gessner MO. Microbial dynamics associated with leaves decomposing in the mainstem and floodplain pond of a large river. *Aquat Microb Ecol* 2002; **28**: 25–36.

Bálint M, Márton O. Proper experimental design requires randomization/balancing of molecular ecology experiments. *bioRxiv* 2017; **87**: 1–19.

Ball DF. Loss-on-ignition as an estimate of organic matter and organic carbon in non-calcareous soils. *J Soil Sci* 1964; **15**: 84–92.

Baranov IV. Biohydrochemical classification of the reservoirs in the European U.S.S.R. 1961; In Tyurin PV (Ed.), *The Storage Lakes of the U.S.S.R. and their Importance for Fishery*. Israel Program for Scientific Translations: Jerusalem, Israel; pp. 139–183.

Barberán A, Bates ST, Casamayor EO, Fierer N. Using network analysis to explore co-occurrence patterns in soil microbial communities. *ISME J* 2012; **6**: 343–351.

Bardgett RD, Van Der Putten WH. Belowground biodiversity and ecosystem functioning. *Nature* 2014; **515**: 505–511.

Bärlocher F, Boddy L. Aquatic fungal ecology - How does it differ from terrestrial? *Fungal Ecol* 2016; **19**: 5–13.

Bastviken D, Cole JJ, Pace ML, Van de-Bogert MC. Fates of methane from different lake habitats: Connecting whole-lake budgets and CH_4 emissions. *J Geophys Res Biogeosciences* 2008; **113**: G02024.

Battin TJ, Kaplan LA, Findlay S, Hopkinson CS, Marti E, Packman AI, Newbold JD, Sabater F. Biophysical controls on organic carbon fluxes in fluvial networks. *Nat Geosci* 2008; **1**:

95–100.

Battin TJ, Luyssaert S, Kaplan LA, Aufdenkampe AK, Richter A, Tranvik LJ. The boundless carbon cycle. *Nat Geosci* 2009; **2**: 598–600.

Becker J, Eisenhauer N, Scheu S, Jousset A. Increasing antagonistic interactions cause bacterial communities to collapse at high diversity. *Ecol Lett* 2012; **15**: 468–474.

Bell T, Newman JA, Silverman BW, Turner SL, Lilley AK. The contribution of species richness and composition to bacterial services. *Nature* 2005; **436**: 1157–1160.

Berggren M, Ström L, Laudon H, Karlsson J, Jonsson A, Giesler R, Bergström AK, Jansson M. Lake secondary production fueled by rapid transfer of low molecular weight organic carbon from terrestrial sources to aquatic consumers. *Ecol Lett* 2010a; **13**: 870–880.

Berggren M, Laudon H, Haei M, Ström L, Jansson M. Efficient aquatic bacterial metabolism of dissolved low-molecular-weight compounds from terrestrial sources. *ISME J* 2010b; **4**: 408–416.

Bertness MD, Callaway R. Positive interactions in communities. *Trends Ecol Evol* 1994; **9**: 187–191.

Besemer K, Singer G, Quince C, Bertuzzo E, Sloan W, Battin TJ. Headwaters are critical reservoirs of microbial diversity for fluvial networks. *Proc R Soc B Biol Sci* 2013; **280**: 20131760.

Bier RL, Bernhardt ES, Boot CM, Graham EB, Hall EK, Lennon JT, Nemergut JR, Osborne BB, Ruiz-González C, Schimel JP, Waldrop MP, Wallenstein MD. Linking microbial community structure and microbial processes: An empirical and conceptual overview. *FEMS Microbiol Ecol* 2015; **91**: fiv113.

Bloom SA. Similarity indices in community studies: potential pitfalls. *Mar Ecol Prog Ser* 1981; **5**: 125–128.

Boisvert-Marsh L, Périé C, De Blois S. Shifting with climate? Evidence for recent changes in tree species distribution at high latitudes. *Ecosphere* 2014; **5**: 1–33.

Bolger AM, Lohse M, Usadel B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* 2014; **30**: 2114–2120.

Borrel G, Jézéquel D, Biderre-Petit C, Morel-Desrosiers N, Morel JP, Peyret P, Fonty G, Lehours AC. Production and consumption of methane in freshwater lake ecosystems. *Res Microbiol* 2011; **162**: 833–847.

Boyer KE, Kertesz JS, Bruno JF. Biodiversity effects on productivity and stability of marine macroalgal communities: The role of environmental context. *Oikos* 2009; **118**: 1062–1072.

Brett M, Muller-Navarra D. The role of highly unsaturated fatty acids in aquatic foodweb processes. *Freshw Biol* 1997; **38**:483–499.

Brett MT, Kainz MJ, Taipale SJ, Seshan H. Phytoplankton, not allochthonous carbon, sustains herbivorous zooplankton production. *Proc Natl Acad Sci USA* 2009; **106**: 21197–21201.

Brett MT, Arhonditsis GB, Chandra S, Kainz MJ. Mass flux calculations show strong allochthonous support of freshwater zooplankton production is unlikely. *PLoS One* 2012; **7**: e39508.

Brown SP, Veach AM, Rigdon-Huss AR, Grond K, Lickteig SK, Lothamer K, Oliver KA, Jumpponen A. Scraping the bottom of the barrel: Are rare high throughput sequences artifacts? *Fungal Ecol* 2015; **13**: 221–225.

Burnham KP, Anderson DR. Model Selection and Multimodel Inference. A Practical Information-Theoretical Approach. 2002; Springer: Heidelberg, Germany.

Cadotte MW, Dinnage R, Tilman D. Phylogenetic diversity promotes ecosystem stability. *Ecology* 2012; **93**: S223–S233.

Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. DADA2: High-resolution sample inference from Illumina amplicon data. *Nat Methods* 2016a; **13**: 581–583.

Callahan BJ, Sankaran K, Fukuyama JA, McMurdie PJ, Holmes SP. Bioconductor Workflow for Microbiome Data Analysis: from raw reads to community analyses. *F1000Res* 2016b; **5**: 1492.

Callahan BJ, McMurdie PJ, Holmes SP. Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *ISME J* 2017; **11**: 2639–2643.

Callaway RM. Positive interactions and interdependence in plant communities. 2007; In *Positive Interactions and Interdependence in Plant Communities*. Springer Netherlands, Dordrecht.

Campbell BJ, Yu L, Heidelberg JF, Kirchman DL. Activity of abundant and rare bacteria in a coastal ocean. *Proc Natl Acad Sci* 2011; **108**: 12776–12781.

Campbell NR, Harmon SA, Narum SR. Genotyping-in-Thousands by sequencing (GT-seq): A cost effective SNP genotyping method based on custom amplicon sequencing. *Mol Ecol Resour* 2015; **15**: 855–867.

Cardinale BJ, Palmer MA, Collins SL. Species diversity enhances ecosystem functioning through interspecific facilitation. *Nature* 2002; **415**: 426–429.

Cardinale BJ, Duffy JE, Gonzalez A, Hooper DU, Perrings C, Venail P, Narwani A, Mace GM, Tilman D, Wardle DA, Kinzig AP, Daily GC, Loreau M, Grace JB, Larigauderie A, Srivastava DS, Naeem. Biodiversity loss and its impact on humanity. *Nature* 2012; **486**: 59–67.

Carlson C, del Giorgio P, Herndl G. Microbes and the Dissipation of Energy and Respiration: From Cells to Ecosystems. *Oceanography* 2007; **20**: 89–100.

Carpenter SR, Cole JJ, Pace ML, Van de Bogert M, Bade DL, Bastviken D, Gille CM, Hodgson JR, Kitchell JF, Kritzberg ES. Ecosystem subsidies: Terrestrial support of aquatic food webs from ^{13}C addition to contrasting lakes. *Ecology* 2005; **86**: 2737–2750.

Chase JM. Stochastic community assembly causes higher biodiversity in more productive environments. *Science* 2010; **328**: 1388–91.

Chen HM, Lifschitz CH. Preparation of fecal samples for assay of volatile fatty acids by gas-liquid chromatography and high-performance liquid chromatography. *Clin Chem* 1989; **35**: 74–76.

Chesson P. General theory of competitive coexistence in spatially-varying environments. *Theor Popul Biol* 2000; **58**: 211–237.

Chung Y, Rabe-Hesketh S, Dorie V, Gelman A, Liu J. A nondegenerate penalized likelihood estimator for variance parameters in multilevel models. *Psychometrika* 2013; **78**: 685-709.

Clark JM, Bottrell SH, Evans CD, Monteith DT, Bartlett R, Rose R, Newton RJ, Chapman PJ. The importance of the relationship between scale and process in understanding long-term DOC dynamics. *Sci Total Environ* 2010; **408**: 2768–2775.

Cleasby IR, Nakagawa S. Neglected biological patterns in the residuals. *Behav Ecol Sociobiol* 2011; **65**: 2361–2372.

Clooney AG, Fouhy F, Sleator RD, O' Driscoll A, Stanton C, Cotter PD, Claesson MJ. Comparing apples and oranges?: Next generation sequencing and its impact on microbiome analysis. *PLoS One* 2016; **11**: 1–16.

Cock PJA, Antao T, Chang JT, Chapman BA, Cox CJ, Dalke A, Friedberg I, Hamelryck T, Kauff F, Wilczynski B, de Hoon MJL. Biopython: Freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics* 2009; **25**: 1422–1423.

Cole JJ, Caraco NF, Kling GW, Kratz TK. Carbon dioxide supersaturation in the surface waters of lakes. *Science* 1994; **265**: 1568–1570.

Cole JJ, Carpenter SR, Pace ML, Van De Bogert MC, Kitchell JL, Hodgson JR. Differential support of lake food webs by three types of terrestrial organic carbon. *Ecol Lett* 2006; **9**: 558–568.

Cole JJ, Prairie YT, Caraco NF, McDowell WH, Tranvik LJ, Striegl RG, DuarteCM, Kortelainen P, Downing JA, Middelburg JJ, Melack J. Plumbing the global carbon cycle: Integrating inland waters into the terrestrial carbon budget. *Ecosystems* 2007; **10**: 171–184.

Creed IF, Bergström AK, Trick CG, Grimm NB, Hessen DO, Karlsson J, Kidd KA, Kritzberg E, McKnight DM, Freeman EC, Senar OE, Andersson A, Ask J, Berggren M, Cherif M, Giesler R, Hotchkiss ER, Kortelainen P, Palta MM, Vrede T, Weyhenmeyer GA. Global change-driven effects on dissolved organic matter composition:Implications for

food webs of northern lakes. *Glob Chang Biol* 2018; **24**: 3692–3714.

Crump BC, Amaral-Zettler LA, Kling GW. Microbial diversity in arctic freshwaters is structured by inoculation of microbes from soils. *ISME J* 2012; **6**: 1629–1639.

Csárdi G, Nepusz T. The igraph software package for complex network research. *InterJournal Complex Syst* 2006; **1695**: 1–9.

Curd EE, Martiny JBH, Li H, Smith TB. Bacterial diversity is positively correlated with soil heterogeneity. *Ecosphere* 2018; **9**: e02079.

de Bello F, Lavorel S, Lavergne S, Albert CH, Boulangeat I, Mazel F, Thuiller W. Hierarchical effects of environmental filters on the functional structure of plant communities: A case study in the French Alps. *Ecography* 2013; **36**: 393–402.

Dean WE, Gorham E. Magnitude and Signifigance of Carbon Burial in Lakes, Reservoirs, and Peatlands. *Geology* 1998; **26**: 535–538.

Del Giorgio PA, Cole JJ, Cimbleris A. Respiration rates in bacteria exceed phytoplankton production in unproductive aquatic systems. *Nature* 1997; **385**: 148–151.

Del Giorgio PA, Duarte CM. Respiration in the open ocean. *Nature* 2002; **420**: 379–384.

Del Giorgio PA. Progress and perspectives in aquatic microbial ecology: Introduction. *Aquat Microb Ecol* 2010; **61**: 219–220.

del Río M, Schütze G, Pretzsch H. Temporal variation of competition and facilitation in mixed species forests in Central Europe. *Plant Biol* 2014; **16**: 166–176.

del Río M, Pretzsch H, Ruíz-Peinado R, Ampoorter E, Annighöfer P, Barbeito I, Bielak K, Brazaitis G, Coll L, Drössler L, Fabrika M, Forrester DI, Heym M, Hurt V, Kurylyak V, Löf M, Lombardi F, Madrickiene E, Matović B, Mohren F, Motta R, den Ouden J, Pach M, Ponette Q, Schütze G, Skrzyszewski J, Sramek V, Sterba H, Stojanović D, Svoboda M, Zlatanov TM, Bravo-Oviedo A. Species interactions increase the temporal stability of community productivity in Pinus sylvestris-Fagus sylvatica mixtures across Europe. *J Ecol* 2017; **105**: 1032–1043.

Delgado-Baquerizo M, Giaramida L, Reich PB, Khachane AN, Hamonts K, Edwards C, Lawton LA, Singh BK. Lack of functional redundancy in the relationship between

microbial diversity and ecosystem functioning. *J Ecol* 2016; **104**: 936–946.

Delgado-Baquerizo M, Reith F, Dennis PG, Hamonts K, Powell JR, Young A, Singh BK, Bissett A. Ecological drivers of soil microbial diversity and soil biological networks in the Southern Hemisphere. *Ecology* 2018a; **99**: 583–596.

Delgado-Baquerizo M, Oliverio AM, Brewer TE, Benavent-González A, Eldridge DJ, Bardgett RD, Maestre FT, Singh BK, Fierer N. A global atlas of the dominant bacteria found in soil. *Science* 2018b; **359**: 320–325.

DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, Huber T, Dalevi D, Hu P, Andersen GL. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl Environ Microb* 2006; **72**: 5069–5072.

de Vries FT, Griffiths RI, Bailey M, Craig H, Girlanda M, Gweon HS, Hallin S, Kaisermann A, Keith AM, Kretzschmar M, Lemanceau P. Soil bacterial networks are less stable under drought than fungal networks. *Nat Commun* 2018; **9**: 3033.

Diaz S, Cabido M, Casanoves F. Plant functional traits and environmental filters at a regional scale. *J Veg Sci* 1998; **9**: 113–122.

Dickie IA, Fukami T, Wilkie JP, Allen RB, Buchanan PK. Do assembly history effects attenuate from species to ecosystem properties? A field test with wood-inhabiting fungi. *Ecol Lett* 2012; **15**: 133–141.

Dillies M-A, Rau A, Aubert J, Hennequet-Antier C, Jeanmougin M, Servant N, Keime C, Marot G, Castel D, Estelle J, Guernec G, Jagla B, Jouneau L, Laloë D, Le Gall C, Schaëffer B, Le Crom S, Guedj M, Jaffrézic F. A comprehensive evaluation of normalization methods for Illumina high-throughput RNA sequencing data analysis. *Brief Bioinform* 2013; **14**: 671–683.

Dillon PJ, Molot LA. Dissolved organic and inorganic carbon mass balances in central Ontario lakes. *Biogeochemistry* 1997; **36**: 29–42.

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**: 1326–1332.

Doolittle WF. Phylogenetic classification and the universal tree. *Science* 1999; **284**: 2124–2128.

Downing AL, Brown BL, Perrin EM, Keitt TH, Leibold MA. Environmental fluctuations induce scale-dependent compensation and increase stability in plankton ecosystems. *Ecology* 2008; **89**: 3204–3214.

Downing AL, Brown BL, Leibold MA. Multiple diversity-stability mechanisms enhance population and community stability in aquatic food webs. *Ecology* 2014; **95**: 173–184.

Drake TW, Raymond PA, Spencer RGM. Terrestrial carbon inputs to inland waters: A current synthesis of estimates and uncertainty. *Limnol Oceanogr Lett* 2017; **3**:132–142.

Duarte CM, Prairie YT. Prevalence of heterotrophy and atmospheric CO2 emissions from aquatic ecosystems. *Ecosystems* 2005; **8**: 862–870.

Edgar RC. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 2010; **26**: 2460–2461.

Einola E, Rantakari M, Kankaala P, Kortelainen P, Ojala A, Pajunen H, Mäkelä S, Arvola L. Carbon pools and fluxes in a chain of five boreal lakes: A dry and wet year comparison. *J Geophys Res Biogeosciences* 2011; **116**: G03009.

Einsele G, Yan J, Hinderer M. Atmospheric carbon burial in modern lake basins and its significance for the global carbon budget. *Glob Planet Change* 2001; **30**: 167–195.

Eisenhauer N, Scheu S, Jousset A. Bacterial diversity stabilizes community productivity. *PLoS One* 2012; **7**: 1–5.

Emilson CE, Kreutzweiser DP, Gunn JM, Mykytczuk NCS. Leaf-litter microbial communities in boreal streams linked to forest and wetland sources of dissolved organic carbon. *Ecosphere* 2017; **8**: e01678.

Emilson EJS, Carson MA, Yakimovich KM, Osterholz H, Dittmar T, Gunn JM, Mykytczuk NCS, Basiliko N, Tanentzap AJ. Climate-driven shifts in sediment chemistry enhance methane production in northern lakes. *Nat Commun* 2018; **9**: 1801.

Evans CD, Monteith DT, Cooper DM. Long-term increases in surface water dissolved organic carbon: Observations, possible causes and environmental impacts. *Environ Pollut* 2005; **137**: 55–71.

Fazi S, Amalfitano S, Pernthaler J, Puddu A. Bacterial communities associated with benthic organic matter in headwater stream microhabitats. *Environ Microbiol* 2005; **7**: 1633–1640.

Fellman JB, Hood E, Spencer RGM. Fluorescence spectroscopy opens new windows into dissolved organic matter dynamics in freshwater ecosystems: A review. *Limnol Oceanogr* 2010; **55**: 2452–2462.

Fierer N, Jackson RB. The diversity and biogeography of soil bacterial communities. *Proc Natl Acad Sci USA* 2006; **103**: 626–631.

Fierer N, Bradford MA, Jackson RB. Toward an ecological classification of soil bacteria. *Ecology* 2007; **88**: 1354–1364.

Fierer N, Nemergut D, Knight R, Craine JM. Changes through time: Integrating microorganisms into the study of succession. *Res Microbiol* 2010; **161**: 635–642.

Fierer N. Embracing the unknown: Disentangling the complexities of the soil microbiome. *Nat Rev Microbiol* 2017; **15**: 579–590.

Findlay S. Aquatic Ecosystems. 2003; In Findlay S, Sinsabaugh RL (Eds.), *Aquatic ecosystems: interactivity of dissolved organic matter*; pp. 363–379.

Finlay BJ, Maberly SC, Cooper JI. Microbial Diversity and Ecosystem Function. *Oikos* 1997; **80**: 209–213.

Finstad G, Helland I, Ugedal O, Hesthagen T, Hessen D. Unimodal response of fish yield to dissolved organic carbon. *Ecol Lett* 2014; **17**: 36–43.

Finstad AG, Andersen T, Larsen S, Tominaga K, Blumentrath S, De Wit HA, Tømmervik H, Hessen DO. From greening to browning: Catchment vegetation development and reduced S-deposition promote organic carbon load on decadal time scales in Nordic lakes. *Sci Rep* 2016; **6**: 31944.

Fitch A, Orland C, Willer D, Emilson EJS, Tanentzap AJ. Feasting on terrestrial organic matter: dining in a dark lake changes microbial decomposition. *Glob Chang Biol* 2018; **24**: 5110-5122.

Flores GE, Henley JB, Fierer N. A Direct PCR Approach to Accelerate Analyses of Human-

Associated Microbial Communities. *PLoS One* 2012; **7**: 1–11.

Flynn DF, Mirotchnick N, Jain M, Palmer MI, Naeem S. Functional and phylogenetic diversity as predictors of biodiversity–ecosystem-function relationships. *Ecology* 2011; **92**: 1573–81.

Freudenstein J V., Broe MB, Folk RA, Sinn BT. Biodiversity and the Species Concept - Lineages are not Enough. *Syst Biol* 2017; **66**: 644–656

Friedline CJ, Franklin RB, McCallister SL, Rivera MC. Microbial community diversity of the eastern Atlantic Ocean reveals geographic differences. *Biogeosciences Discuss* 2012; **9**: 109–150.

Fuhrman JA, Steele JA. Community structure of marine bacterioplankton: Patterns, networks, and relationships to function. *Aquat Microb Ecol* 2008; **53**: 69–81.

Fukami T, Dickie IA, Paula Wilkie J, Paulus BC, Park D, Roberts A, Buchanan PK, Allen RB. Assembly history dictates ecosystem functioning: Evidence from wood decomposer communities. *Ecol Lett* 2010; **13**: 675–684.

Fukami T. Historical Contingency in Community Assembly: Integrating Niches, Species Pools, and Priority Effects. *Annu Rev Ecol Evol Syst* 2015; **46**: 1–23.

Fukami T, Morin PJ. Productivity-biodiversity relationships depend on the history of community assembly. *Nature* 2003; **424**: 423–426.

Garcia JL, Patel BK, Ollivier B. Taxonomic, phylogenetic, and ecological diversity of methanogenic Archaea. *Anaerobe* 2000; **6**: 205–226.

Gardes M, Bruns TD. ITS primers with enhanced specificity for basidiomycetes - application to the identification of mycorrhizae and rusts. *Mol Ecol* 1993; **2**: 113–118.

Girvan MS, Campbell CD, Killham K, Prosser JI, Glover LA. Bacterial diversity promotes community stability and functional resilience after perturbation. *Environ Microbiol* 2005; **7**: 301–313.

Golchin A, Oades JM, Skjemstad JO, Clarke P. Soil structure and carbon cycling. *Soil Res* 1994; **32**: 1043–1068.

Gotelli NJ, McCabe DJ. Species co-occurrence: a meta-analysis of J. M. Diamond's assembly

rules model. *Ecology* 2002; **83**: 2091–2096.

Grace JB, Schoolmaster DR, Guntenspergen GR, Little AM, Mitchell BR, Miller KM, Schweiger EW. Guidelines for a graph-theoretic implementation of structural equation modeling. *Ecosphere* 2012; **3**: 1–44.

Graham EB, Knelman JE, Schindlbacher A, Siciliano S, Breulmann M, Yannarell A, Beman JM, Abell G, Philippot L, Prosser J, Foulquier A, Yuste JC, Glanville HC, Jones DL, Angel R, Salminen J, Newton RJ, Bürgmann H, Ingram LJ, Hamer U, Siljanen HMP, Peltoniemi K, Potthast K, Bañeras L, Hartmann M, Banerjee S, Yu RQ, Nogaro G, Richter A, Koranda M, Castle SC, Goberna M, Song B, Chatterjee A, Nunes OC, Lopes AR,Cao Y, Kaisermann A, Hallin S, Strickland MS, Garcia-Pausas J, Barba J, Kang H, Isobe K, Papaspyrou S, Pastorelli R, Lagomarsino A, Lindström ES, Basiliko N, Nemergut DR. Microbes as engines of ecosystem function: When does community structure enhance predictions of ecosystem processes? *Front Microbiol* 2016; **7**: 214.

Grenfell BT, Wilson K, Finkenstädt BF, Coulson TN, Murray S, Albon SD, Pemberton JM, Clutton-Brock TH, Crawley MJ. Noise and determinism in synchronized sheep dynamics. *Nature* 1998; **394**: 674–677.

Grman E, Lau JA, Schoolmaster DR, Gross KL. Mechanisms contributing to stability in ecosystem function depend on the environmental context. *Ecol Lett* 2010; **13**: 1400–1410.

Gross K, Cardinale BJ, Fox JW, Gonzalez A, Loreau M, Wayne Polley H, Reich PB, Ruijven J. Species Richness and the Temporal Stability of Biomass Production: A New Analysis of Recent Biodiversity Experiments. *Am Nat* 2014; **183**: 1–12.

Grossart HP, Rojas-Jimenez K. Aquatic fungi: Targeting the forgotten in microbial ecology. *Curr Opin Microbiol* 2016; **31**: 140–145.

Grundmann GL. Geostatistical analysis of the distribution of NH4 and NO^{2-} oxidizing bacteria and serotypes at the millimeter scale along a soil transect. *FEMS Microbiol Ecol* 2000; **34**: 57–62.

Gu Z, Gu L, Eils R, Schlesner M, Brors B. Circlize implements and enhances circular

visualization in R. *Bioinformatics* 2014; **30**: 2811–2812.

Gudasz C, Ruppenthal M, Kalbitz K, Cerli C, Fiedler S, Oelmann Y, Andersson A, Karlsson J. Contributions of terrestrial organic carbon to northern lake sediments. *Limnol Oceanogr Lett* 2017; **2**:218–227.

Guillemette F, del Giorgio PA. Reconstructing the various facets of dissolved organic carbon bioavailability in freshwater ecosystems. *Limnol Oceanogr* 2011; **56**: 734–748.

Guillemette F, Leigh Mccallister S, Del Giorgio PA. Selective consumption and metabolic allocation of terrestrial and algal carbon determine allochthony in lake bacteria. *ISME J* 2016; **10**: 1373–1382.

Gulis V, Suberkropp K. Leaf litter decomposition and microbial activity in nutrient-enriched and unaltered reaches of a headwater stream. *Freshw Biol* 2003; **48**: 123–134.

Haegeman B, Hamelin J, Moriarty J, Neal P, Dushoff J, Weitz JS. Robust estimation of microbial diversity in theory and in practice. *ISME J* 2013; **7**: 1092–1101.

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

Hanson PC, Carpenter SR, Armstrong DE, Stanley EH, Kratz TK. Lake Dissolved Inorganic Carbon and Dissolved Oxygen : Changing Drivers from Days to Decades. *Ecol Monogr* 2006; **76**: 343–363.

Hanson PC, Pace ML, Carpenter SR, Cole JJ, Stanley EH. Integrating Landscape Carbon Cycling: Research Needs for Resolving Organic Carbon Budgets of Lakes. *Ecosystems* 2015; **18**: 363–375.

Hautier Y, Seabloom EW, Borer ET, Adler PB, Harpole WS, Hillebrand H, Lind EM, MacDougall AS, Stevens CJ, Bakker JD, Buckley YM, Chu C, Collins SL, Daleo P, Damschen EI, Davies KF, Fay PA, Firn J, Gruner DS, Jin VL, Klein JA, Knops JMH, La Pierre KJ, Li W, McCulley RL, Melbourne BA, Moore JL, O'Halloran LR, Prober SM, Risch AC, Sankaran M, Schuetz M, Hector A. Eutrophication weakens stabilizing effects of diversity in natural grasslands. *Nature* 2014; **508**: 521–525.

Heathcote AJ, Anderson NJ, Prairie YT, Engstrom DR, Del Giorgio PA. Large increases in

carbon burial in northern lakes during the Anthropocene. *Nat Commun* 2015; **6**: 1–6.

Hector A, Hautier Y, Saner P, Wacker L, Bagchi R, Joshi J, Scherer-Lorenzen M, Spehn EM, Bazeley-White E, Weilenmann M, Caldeira MC, Dimitrakopoulos PG, Finn JA, Huss-Danell K, Jumpponen A, Mulder CPH, Palmborg C, Pereira JS, Siamantziouras ASD, Terry AC, Troumbis AY, Schmid B, Loreau M. General stabilizing effects of plant diversity on grassland productivity through population asynchrony and overyielding. *Ecology* 2010; **91**: 2213–2220.

Hendershot JN, Read QD, Henning JA, Sanders NJ, Classen AT. Consistently inconsistent drivers of microbial diversity and abundance at macroecological scales. *Ecology* 2017; **98**: 1757–1763.

Hessen DO. Dissolved organic carbon in a humic lake: effects on bacterial production and respiration. *Hydrobiologia* 1992; **229**: 115–123.

Hessen DO. Food webs and carbon cycling in humic lakes. 1998; In Hessen DO, Tanvik LJ (Eds.), *Aquatic humic substances*. Springer, Berlin, Heidelberg, pp. 285–315.

Hewson I, Steele JA, Capone DG, Fuhrman JA. Temporal and spatial scales of variation in bacterioplankton assemblages of oligotrophic surface waters. *Mar Ecol Prog Ser* 2006; **311**: 67–77.

HilleRisLambers J, Adler PB, Harpole WS, Levine JM, Mayfield MM. Rethinking Community Assembly through the Lens of Coexistence Theory. *Annu Rev Ecol Evol Syst* 2012; **43**: 227–248.

Hooper DU, Solan M, Symstad AJ, Diaz S, Gessner MO, Buchmann N, Degrange V, Grime JP, Hulot FD, Mermillod-Blondin F, Roy J, Spehn EM, Van Peer L. Species diversity, functional diversity, and ecosystem functioning. *Biodiversity and ecosystem functioning: Syntheses and perspectives* 2002; **17**: 195–208.

Hooper DU, Dukes JS. Overyielding among plant functional groups in a long-term experiment. *Ecol Lett* 2004; **7**: 95–105.

Hooper DU, Adair EC, Cardinale BJ, Byrnes JEK, Hungate B a., Matulich KL, Gonzalez A, Duffy JE, Gamfeldt L, O'Connor MI. A global synthesis reveals biodiversity loss as a

major driver of ecosystem change. *Nature* 2012; **486**: 105–108.

Houlahan JE, Currie DJ, Cottenie K, Cumming GS, Ernest SKM, Findlay CS, Fuhlendorf SD, Gaedke U, Legendre P, Magnuson JJ, McArdle BH, Muldavin EH, Noble D, Russell R, Stevens RD, Willis TJ, Woiwod IP, Wondzell SM. Compensatory dynamics are rare in natural ecological communities. *Proc Natl Acad Sci* 2007; **104**: 3273–3277.

Hugerth LW, Andersson AF. Analysing microbial community composition through amplicon sequencing: From sampling to hypothesis testing. *Front Microbiol* 2017; **8**: 1561.

Hunting ER, Vijver MG, van der Geest HG, Mulder C, Kraak MHS, Breure AM, Admiraal W. Resource niche overlap promotes stability of bacterial community metabolism in experimental microcosms. *Front Microbiol* 2015; **6**: 1–7.

Huryn AD, Butz Huryn VM, Arbuckle CJ, Tsomides L. Catchment land-use, macroinvertebrates and detritus processing in headwater streams: Taxonomic richness versus function. *Freshw Biol* 2002; **47**: 401–415.

Isbell F, Polley HW, Wilsey BJ. Biodiversity, productivity and the temporal stability of productivity : patterns and processes. *Ecol Lett* 2009; **12**: 443–451.

Ives AR, Gross K, Klug JL. Stability and variability in competitive communities. *Science* 1999; **286**: 542–544.

Jackson CR, Churchill PF, Roden EE. Successional Changes in Bacterial Assemblage Structure during Epilithic Biofilm Development. *Ecology* 2014; **82**: 555–566.

Jansson M, Bergström A-K, Blomqvist P, Drakare S. Allochthonous organic carbon and phytoplankton/bacterioplankton produciton relationships in lakes. *Ecology* 2000; **81**: 3250–3255.

Jansson M, Persson L, De Roos AM, Jones RI, Tranvik LJ. Terrestrial carbon and intraspecific size-variation shape lake ecosystems. *Trends Ecol Evol* 2007; **22**: 316–322.

Jiang L. Negative selection effects suppress relationships between bacterial diversity and ecosystem functioning. *Ecology* 2007; **88**: 1075–1085.

Jones P, Binns D, Chang HY, Fraser M, Li W, McAnulla C, McWilliam H, Maslen J, Mitchell A, Nuka G, Pesseat S, Quinn AF, Sangrador-Vegas A, Scheremetjew M, Yong SY,

Lopez R, Hunter S. InterProScan 5: genome-scale protein function classification. *Bioinformatics* 2014; **30**: 1236–1240.

Jones SE, Lennon JT, Kellogg WK. Dormancy contributes to the maintenance of microbial diversity. *PNAS* 2010; **107**: 5881–5886.

Jones SE, Newton RJ, McMahon KD. Evidence for structuring of bacterial community composition by organic carbon source in temperate lakes. *Environ Microbiol* 2009; **11**: 2463–2472.

Jones SE, Solomon CT, Weidel BC. Subsidy or Subtraction: How Do Terrestrial Inputs Influence Consumer Production in Lakes? *Freshw Rev* 2012; **5**: 37–49.

Jones SE, Lennon JT. A test of the subsidy-stability hypothesis: The effects of terrestrial carbon in aquatic ecosystems. *Ecology* 2015; **96**: 1550–1560.

Jonsson A, Meili M, Bergström A-K, Jansson M. Whole-lake mineralization of allochthonous and autochthonous organic carbon in a large humic lake (Örträsket, N. Sweden). *Limnol Oceanogr* 2001; **46**: 1691–1700.

Jousset A, Bienhold C, Chatzinotas A, Gallien L, Gobet A, Kurm V, Küsel K, Rillig MC, Rivett DW, Salles JF, van der Heijden MGA, Youssef NH, Zhang X, Wei Z, Hol WHG. Where less may be more: How the rare biosphere pulls ecosystems strings. *ISME J* 2017; **11**: 853–862.

Jucker T, Coomes DA. Comment on 'Plant species richness and ecosystem multifunctionality in global drylands'. *Science* 2012; **337**: 155–155.

Jucker T, Bouriaud O, Avacaritei D, Coomes DA. Stabilizing effects of diversity on aboveground wood production in forest ecosystems: linking patterns and processes. *Ecol Lett* 2014; **17**: 1560–1569.

Judd KE, Crump BC, Kling GW. Variation in dissolved organic matter controls bacterial production and community composition. *Ecology* 2006; **87**: 2068–2079.

Judd KE, Crump BC, Kling GW. Bacterial responses in activity and community composition to photo-oxidation of dissolved organic matter from soil and surface waters. *Aquat Sci* 2007; **69**: 96–107.

Kara EL, Hanson PC, Hu YH, Winslow L, McMahon KD. A decade of seasonal dynamics and co-occurrences within freshwater bacterioplankton communities from eutrophic Lake Mendota, WI, USA. *ISME J* 2013; **7**: 680–684.

Karlsson J, Byström P, Ask J, Ask P, Persson L, Jansson M. Light limitation of nutrient-poor lake ecosystems. *Nature* 2009; **460**: 506–509.

Karlsson J, Berggren M, Ask J, Byström P, Jonsson A, Laudon H, Jansson M. Terrestrial organic matter support of lake food webs: Evidence from lake metabolism and stable hydrogen isotopes of consumers. *Limnol Oceanogr* 2012; **57**: 1042–1048.

Kellerman AM, Kothawala DN, Dittmar T, Tranvik LJ. Persistence of dissolved organic matter in lakes related to its molecular characteristics. *Nat Geosci* 2015; **8**: 454–457.

Kent AD, Yannarell AC, Rusak JA, Triplett EW, McMahon KD. Synchrony in aquatic microbial community dynamics. *ISME J* 2007; **1**: 38–47.

Kirchman DL. The ecology of Cytophaga-Flavobacteria in aquatic environments. *FEMS Microbiol Ecol* 2002; **39**: 91–100.

Kirchman DL. Degradation of organic matter. 2002; In Kirchman DL (Ed.), *Processes in Microbial Ecology*. Oxford University Press: Oxford, UK; pp. 113-133.

Kirk TK, Farrell RL. Enzymatic" combustion": the microbial degradation of lignin. *Annu Rev Microbiol* 1987; **41**: 465–501.

Kirtman B, Power SB, Adedoyin AJ, Boer GJ, Bojariu R, Camilloni I, Doblas-Reyes F, Fiore AM, Kimoto M, Meehl G, Prather M, Sarr A, Schär C, Sutton R, van Oldenborgh GJ, Vecchi G, Wang HJ. Near-term climate change: Projections and predictability. 2013; In Stocker TF, Qin D, Plattner GK, Tignor M, Allen SK, Boschung J, Nauels A, Xia Y, Bex V, Midgley PM (Eds.), *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge, UK and New York, NY: Cambridge University Press; pp. 953–1028.

Klindworth A, Pruesse E, Schweer T, Peplies J, Quast C, Horn M, Glöckner FO. Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation

sequencing-based diversity studies. *Nucleic Acids Res* 2013; **41**: 1–11.

Knight R, Vrbanac A, Taylor BC, Aksenov A, Callewaert C, Debelius J, Gonzalez A, Kosciolek T, McCall LI, McDonald D, Melnik AV, Morton JT, Navas J, Quinn RA, Sanders JG, Swafford AD, Thompson LR, Tripathi A, Xu ZZ, Zaneveld JR, Zhu Q, Caporaso JG, Dorrestein PC. Best practices for analysing microbiomes. *Nat Rev Microbiol* 2018; **16**: 410–422.

Kõljalg U, Nilsson RH, Abarenkov K, Tedersoo L, Taylor AFS, Bahram M, Bates ST, Bruns TD, Bengtsson-Palme J, Callaghan TM, Douglas B, Drenkhan T, Eberhardt U, Dueñas M, Grebenc T, Griffith GW, Hartmann M, Kirk PM, Kohout P, Larsson E, Lindahl BD, Lücking R, Martín MP, Matheny PB, Nguyen NH, Niskanen T, Oja J, Peay KG, Peintner U, Peterson M, Põldmaa K, Saag L, Saar I, Schüßler A, Scott JA, Senés C,Smith ME, Suija A, Taylor DL, Telleria MT, Weiss M, Larsson KH. Towards a unified paradigm for sequence-based identification of fungi. *Mol Ecol* 2013; **22**: 5271–5277.

Koranda M, Kaiser C, Fuchslueger L, Kitzler B, Sessitsch A, Zechmeister-Boltenstern S, Richter A. Fungal and bacterial utilization of organic substrates depends on substrate complexity and N availability. *FEMS Microbiol Ecol* 2014; **87**: 142–152.

Kraft NJB, Adler PB, Godoy O, James EC, Fuller S, Levine JM. Community assembly, coexistence and the environmental filtering metaphor. *Funct Ecol* 2015; **29**: 592–599.

Krause S, Le Roux X, Niklaus PA, van Bodegom PM, Lennon T. JT, Bertilsson S, Grossart HP, Philippot L, Bodelier PLE. Trait-based approaches for understanding microbial biodiversity and ecosystem functioning. *Front Microbiol* 2014; **5**: 251

Kuehn KA. Lentic and lotic habitats as templets for fungal communities: Traits, adaptations, and their significance to litter decomposition within freshwater ecosystems. *Fungal Ecol* 2016; **19**: 135–154.

Kuiters AT, Sarink HM. Leaching of phenolic compounds from leaf and needle litter of several deciduous and coniferous trees. *Soil Biol Biochem* 1986; **18**: 475–480.

Laforest-Lapointe I, Paquette A, Messier C, Kembel SW. Leaf bacterial diversity mediates plant

diversity and ecosystem function relationships. *Nature* 2017; **546**: 145–147.

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

Lapierre JF, Guillemette F, Berggren M, Del Giorgio PA. Increases in terrestrially derived carbon stimulate organic carbon processing and CO 2 emissions in boreal aquatic ecosystems. *Nat Commun* 2013; **4**: 734–748.

Lau DCP, Sundh I, Vrede T, Pickova J, Goedkoop W. Autochthonous resources are the main driver of consumer production in dystrophic boreal lakes. *Ecology* 2014; **95**: 1506–1519.

Lavonen EE, Kothawala DN, Tranvik LJ, Gonsior M, Schmitt-Kopplin P, Köhler SJ. Tracking changes in the optical properties and molecular composition of dissolved organic matter during drinking water production. *Water Res* 2015; **85**: 286–294.

Lear G, Bellamy J, Case BS, Lee JE, Buckley HL. Fine-scale spatial patterns in bacterial community composition and function within freshwater ponds. *ISME J* 2014; **8**: 1715–1726.

Lefébure R, Degerman R, Andersson A, Larsson S, Eriksson LO, Båmstedt U, Byström P. Impacts of elevated terrestrial nutrient loads and temperature on pelagic food-web efficiency and fish production. *Glob Chang Biol* 2013; **19**: 1358–1372.

Leibold MA, Mcpeek MA. Coexistence of the Niche and Neutral Perspectives in Community Ecology. *Ecology* 2016; **87**: 1399–1410.

Lewandowski D, Kurowicka D, Joe H. Generating random correlation matrices based on vines and extended onion method. *J Multivariate Anal* 2009; **100**: 1989–2001.

Lindström ES, Langenheder S. Local and regional factors influencing bacterial community assembly. *Environ Microbiol Rep* 2012; **4**: 1–9.

Loreau M, de Mazancourt C. Species Synchrony and Its Drivers: Neutral and Nonneutral Community Dynamics in Fluctuating Environments. *Am Nat* 2008; **172**: E48–E66.

Loreau M, de Mazancourt C. Biodiversity and ecosystem stability: A synthesis of underlying mechanisms. *Ecol Lett* 2013; **16**: 106–115.

Loreau M, Hector A. Partitioning selection and complementarity in biodiversity experiments.

Nature 2001; **412**: 72–76.

Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 2014; **15**: 550.

Lovley DR, Klug MJ. Sulfate Reducers Can Outcompete Methanogens at Concentrations Sulfate Reducers Can Outcompete Methanogens Sulfate Concentrationst at Freshwater. *Appl Environ Microbiol* 1983; **45**: 187–194.

Lyautey E, Jackson CR, Cayrou J, Rols JL, Garabétian F. Bacterial community succession in natural river biofilm assemblages. *Microb Ecol* 2005; **50**: 589–601.

Lynch MDJ, Neufeld JD. Ecology and exploration of the rare biosphere. *Nat Rev Microbiol* 2015; **13**: 217–229.

Maestre FT, Callaway RM, Valladares F, Lortie CJ. Refining the stress-gradient hypothesis for competition and facilitation in plant communities. *J Ecol* 2009; **97**: 199–205.

Magurran AE, Baillie SR, Buckland ST, Dick JM, Elston DA, Scott EM, Smith RI, Somerfield PJ, Studeny AC, Watt A. Long-term datasets in biodiversity research and monitoring: assessing change in ecological communities through time. *Trends Ecol Evol* 2010; **25**: 574–582.

Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* 2011; **17**: 10-12.

Martiny AC, Jørgensen TM, Albrechtsen HJ, Arvin E, Molin S. Long-Term Succession of Structure and Diversity of a Biofilm Formed in a Model Drinking Water Distribution System. *Appl Environ Microbiol* 2003; **69**: 6899–6907.

Martiny AC, Treseder K, Pusch G. Phylogenetic conservatism of functional traits in microorganisms. *ISME J* 2013; **7**: 830–838.

Martiny JBH, Bohannan BJM, Brown JH, Colwell RK, Fuhrman JA, Green JL, Horner-Devine MC, Kane M , Adams Krumins J, Kuske CR, Morin PJ, Naeem S, Øvreås L, Reysenbach AL, Smith VH, Staley JT Microbial biogeography: Putting microorganisms on the map. *Nat Rev Microbiol* 2006; **4**: 102–112.

Martiny JBH, Jones SE, Lennon JT, Adam C. Microbiomes in light of traits : a phylogenetic

perspective. 2015; *Science* **350**: aac9323.

Martiny JBH, Martiny AC, Weihe C, Lu Y, Berlemont R, Brodie EL, Goulden ML, Treseder KK, Allison SD. Microbial legacies alter decomposition in response to simulated global change. *ISME J* 2017; **11**: 490–499.

McCann KS. The diversity–stability debate. *Nature* 2000; **405**: 228–233.

McGrady-Steed J, Harris PM, Morin PJ. Biodiversity regulates ecosystem predictability. *Nature* 1997; **390**: 162–165.

McKenney DW, Pedlar JH, Lawrence K, Papadopol P, Campbell K, Hutchinson MF. Change and evolution in the plant hardiness zones of Canada. *Bioscience* 2014; **64**: 341–350

McMurdie PJ, Holmes S. Waste Not, Want Not: Why Rarefying Microbiome Data Is Inadmissible. *PLoS Comput Biol* 2014; **10**: e1003531.

Micheli F, Cottingham KL, Bascompte J, Bjornstad ON, Eckert GL, Fischer JM, Keitt TH, Kendall BE, Klug JL, Rusak JA. The Dual Nature of Community Variability. *Oikos* 1999; **85**: 161-169.

Miller DN, Bryant JE, Madsen EL, Ghiorse WC. Evaluation and optimization of DNA extraction and purification procedures for soil and sediment samples. *Appl Environ Microbiol* 1999; **65**: 4715–4724.

Mitchell A, Bucchini F, Cochrane G, Denise H, Ten Hoopen P, Fraser M, Pesseat S, Potter S, Scheremetjew M, Sterk P, Finn RD. EBI metagenomics in 2016 - An expanding and evolving resource for the analysis and archiving of metagenomic data. *Nucleic Acids Res* 2016; **44**: 595–603.

Morel FM, Hering JG. Principles and applications of aquatic chemistry. 1993; John Wiley & Sons: NY, USA.

Morriën E, Hannula SE, Snoek LB, Helmsing NR, Zweers H, de Hollander M, Soto RL, Bouffaud ML, Buée M, Dimmers W, Duyts H, Geisen S, Girlanda M, Rob I. Griffiths RI, Jørgensen HB, Jensen J, Plassart P, Redecker D, Schmelz RM, Schmidt O, Thomson BC, Tisserant E, Uroz S, Winding A, Bailey MJ, Bonkowski M, Faber JH,Martin F, Lemanceau P, de Boer W, van Veen JA, Wim H, van der Putten WH. Soil networks

become more connected and take up more carbon as nature restoration progresses. *Nat Commun* 2017; **8**: 14349.

Mulder CP, Uliassi DD, Doak DF. Physical stress and diversity-productivity relationships: the role of positive interactions. *Proc Natl Acad Sci USA* 2001; **98**: 6704–6708.

Muscarella ME, Jones SE, Lennon JT. Species sorting along a subsidy gradient alters bacterial community stability. *Ecology* 2016; **97**: 2034–2043.

Naeem S. Disentangling the Impacts of Diversity on Ecosystem Functioning in Combinatorial Experiments. *Ecology* 2002; **83**: 2925–2935.

Natelhoffer KJ, Fry B. Controls on natural nitrogen-15 and carbon-13 abundances in forest soil organic matter. *Soil Sci Soc Am J* 1988; **52**: 1633–1640.

Nelson CE, Sadro S, Melack JM. Contrasting the influences of stream inputs and landscape position on bacterioplankton community structure and dissolved organic matter composition in high- elevation lake chains. *Limnol Ocean* 2009; **54**: 1292–1305.

Nemergut DR, Anderson SP, Cleveland CC, Martin AP, Miller AE, Seimon A, Schmidt SK. Microbial community succession in an unvegetated, recently deglaciated soil. *Microb Ecol* 2007; **53**: 110–122.

Nemergut DR, Schmidt SK, Fukami T, O'Neill SP, Bilinski TM, Stanish LF, Knelman JE, Darcy JL, Lynch RC, Wickey P, Ferrenberg S. Patterns and processes of microbial community assembly. *Microbiol Mol Biol Rev* 2013; **77**: 342–56.

Newman MEJ. The Structure and Function of Complex Networks. *SIAM Rev* 2003; **45**: 167–256.

Nielsen UN, Ayres E, Wall DH, Bardgett RD. Soil biodiversity and carbon cycling: A review and synthesis of studies examining diversity-function relationships. *Eur J Soil Sci* 2011; **62**: 105–116.

Nilakanta H, Drews KL, Firrell S, Foulkes MA, Jablonski KA. A review of software for analyzing molecular sequences. *BMC Res Notes* 2014; **7**:830.

Nilsson RH, Kristiansson E, Ryberg M, Hallenberg N, Larsson KH. Intraspecific ITS variability in the Kingdom Fungi as expressed in the international sequence databases and its

implications for molecular species identification. *Evol Bioinform* 2008; **2008**: 193–201.

Nino-Garcia JP, Ruiz-Gonzalez C, del Giorgio PA. Landscape-scale spatial abundance distributions discriminate core from random components of boreal lake bacterioplankton. *Ecol Lett* 2016; **19**: 1506–1515.

O'Donnell JA, Aiken GR, Walvoord MA, Butler KD. Dissolved organic matter composition of winter flow in the Yukon River basin: Implications of permafrost thaw and increased groundwater discharge. *Global Biogeochem Cycles* 2012; **26**: GB0306.

Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PR, O'Hara RB, Simpson GL, Solymos P, Stevens MHH, Wagner H. vegan: Community ecology package. 2013.

Ontario Ministry of the Environment [OMOE]. Handbook of analytical methods for environmental samples. 1983; Ontario Ministry of the Environment, Rexdale, Canada.

Orland C, Emilson EJS, Basiliko N, Mykytczuk NCS, Gunn JM, Tanentzap AJ. Microbiome functioning depends on individual and interactive effects of the environment and community structure. *ISME J* 2018; in press, https://doi.org/10.1038/s41396-018-0230-x.

Pace ML, Cole JJ, Carpenter SR. Whole-lake carbon-13 additions reveal terrestrial support of aquatic food webs. *Nature* 2004; **427**: 240–243.

Paerl HW, Dyble J, Twomey L, Pinckney JL, Nelson J, Kerkhof L. Characterizing man-made and natural modifications of microbial diversity and activity in coastal ecosystems. *Antonie van Leeuwenhoek, Int J Gen Mol Microbiol* 2002; **81**: 487–507.

Pala C, Molari M, Nizzoli D, Bartoli M, Viaroli P, Manini E. Environmental Drivers Controlling Bacterial and Archaeal Abundance in the Sediments of a Mediterranean Lagoon Ecosystem. *Curr Microbiol* 2018; **75**: 1147–1155.

Patin N V, Kunin V, Lidström U, Ashby MN. Effects of OTU clustering and PCR artifacts on microbial diversity estimates. *Microb Ecol* 2013; **65**: 709–719.

Pedrós-Alió C. Dipping into the rare biosphere. *Science* 2007; **315**: 192–193.

Peralta AL, Johnston ER, Matthews JW, Kent AD. Abiotic correlates of microbial community structure and nitrogen cycling functions vary within wetlands. *Freshw Sci* 2016; **35**: 573–588.

Peura S, Eiler A, Hiltunen M, Nykänen H, Tiirola M, Jones RI. Bacterial and phytoplankton responses to nutrient amendments in a boreal lake differ according to season and to taxonomic resolution. *PLoS One* 2012; **7**: e38552.

Pion M, Spangenberg JE, Simon A, Bindschedler S, Flury C, Chatelain A, Bshary R, Job D, Junier P. Bacterial farming by the fungus Morchella crassipes. *Proc R Soc B Biol Sci* 2013; **280**: 20132242.

Pollock J, Glendinning L, Wisedchanwet T, Watson M. The madness of microbiome: Attempting to find consensus “best practice” for 16S microbiome studies. *Appl Environ Microbiol* 2018; **84**: e02627-02617.

Pomeroy LR. The Ocean’s Food Web, A Changing Paradigm. *Bioscience* 1974; **24**: 499–504.

Ponomarova O, Patil KR. Metabolic interactions in microbial communities: Untangling the Gordian knot. *Curr Opin Microbiol* 2015; **27**: 37–44.

Popp D, von Gillhaussen P, Weidlich EWA, Sträuber H, Harms H, Temperton VM. Methane yield of biomass from extensive grassland is affected by compositional changes induced by order of arrival. *GCB Bioenergy* 2017; **9**: 1555–1562.

Powell JR, Welsh A, Hallin S, Allison SD. Microbial functional diversity enhances predictive models linking environmental parameters to ecosystem properties. *Ecology* 2015; **96**: 1985–1993.

Ptacnik R, Solimini AG, Andersen T, Tamminen T, Brettum P, Lepisto L, Willén E, Rekolainen S. Diversity predicts stability and resource use efficiency in natural phytoplankton communities. *Proc Natl Acad Sci* 2008; **105**: 5134–5138.

Purvis A, Hector A. Getting the measure of biodiversity. *Nature* 2000; **405**: 212-219.

Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO. The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Res* 2013; **41**: 590–596.

Ramus AP, Long ZT. Producer diversity enhances consumer stability in a benthic marine community. *J Ecol* 2016; **104**: 572–579.

Ranjan R, Rani A, Metwally A, McGee HS, Perkins DL. Analysis of the microbiome:

Advantages of whole genome shotgun versus 16S amplicon sequencing. *Biochem Biophys Res Commun* 2016; **469**: 967–977.

Ravinet M, Syväranta J, Jones RI, Grey J. A trophic pathway from biogenic methane supports fish biomass in a temperate lake ecosystem. *Oikos* 2010; **119**: 409–416.

Reche I, Pace ML, Cole JJ. Interactions of photobleaching and inorganic nutrients in determining bacterial growth on colored dissolved organic carbon. *Microb Ecol* 1998; **36**: 270–280.

Redford AJ, Fierer N. Bacterial succession on the leaf surface: A novel system for studying successional dynamics. *Microb Ecol* 2009; **58**: 189–198.

Reed HE, Martiny JBH. Testing the functional significance of microbial composition in natural communities. *FEMS Microbiol Ecol* 2007; **62**: 161–170.

Reed HE, Martiny JBH. Microbial composition affects the functioning of estuarine sediments. *ISME J* 2013; **7**: 868–879.

Rho M, Tang H, Ye Y. FragGeneScan: predicting genes in short and error-prone reads. *Nucleic Acids Res* 2010; **38**: 191.

Richardson JS. Seasonal food limitation of detritivores in a montane stream: an experimental test. *Ecology* 1991; **72**: 873–887.

Rocca JD, Hall EK, Lennon JT, Evans SE, Waldrop MP, Cotner JB, Nemergut DR, Graham EB, Wallenstein MD. Relationships between protein-encoding gene abundance and corresponding process are commonly assumed yet rarely observed. *ISME J* 2015; **9**: 1693–1699.

Roehm CL, Prairie YT, Del Giorgio PA. The pCO2 dynamics in lakes in the boreal region of northern Québec, Canada. *Global Biogeochem Cycles* 2009; **23**: 1–9.

Roger F, Bertilsson S, Langenheder S, Osman OA, Gamfeldt L. Effects of multiple dimensions of bacterial diversity on functioning, stability and multifunctionality. *Ecology* 2016; **97**: 2716–2728.

Romanuk TN, Vogt RJ, Young A, Tuck C, Carscallen MW. Maintenance of positive diversity-stability relations along a gradient of environmental stress. *PLoS One* 2010; **5**: e10378.

Roscher C, Weigelt A, Proulx R, Marquard E, Schumacher J, Weisser WW, Schmid B. Identifying population- and community-level mechanisms of diversity-stability relationships in experimental grasslands. *J Ecol* 2011; **99**: 1460–1469.

Ruiz-González C, Niño-García JP, Lapierre J-F, del Giorgio PA. The quality of organic matter shapes the functional biogeography of bacterioplankton across boreal freshwater ecosystems. *Glob Ecol Biogeogr* 2015a; **24**: 1487–1498.

Ruiz-González C, Nino-Garcia JP, del Giorgio PA. Terrestrial origin of bacterial communities in complex boreal freshwater networks. *Ecol Lett* 2015b; **18**: 1198–1206.

Ruiz-gonzález C, Archambault E, Laforest-lapointe I, Giorgio PA, Steven W, Messier C, Nock CA, Beisner BE. Soils associated to different tree communities do not elicit predictable responses in lake bacterial community structure and function. *FEMS Microbiol Ecol* 2018; **94**: fiy115.

Rummens K, De Meester L, Souffreau C. Inoculation history affects community composition in experimental freshwater bacterioplankton communities. *Environ Microbiol* 2018; **20**: 1120–1133.

Salter SJ, Cox MJ, Turek EM, Calus ST, Cookson WO, Moffatt MF, Turner P, Parkhill J, J Loman NJ, Walker AW. Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biol* 2014; **12**: 87.

Sankaran M, McNaughton SJ. Determinants of biodiversity regulate compositional stability of communities. *Nature* 1999; **401**: 691–693.

Schallenberg M, Hamilton DP, Hicks AS, Robertson HA, Scarsbrook M, Robertson B, Wilson K, Whaanga D, Jones HFE, Hamill K. Multiple lines of evidence determine robust nutrient load limits required to safeguard a threatened lake/lagoon system. *New Zeal J Mar Freshw Res* 2017; **51**: 78–95.

Schallenberg M, Kalff J. The ecology of sediment bacteria in lakes and comparisons with other aquatic ecosystems. *Ecology* 1993; **74**: 919–934.

Schindler DW, Lee PG. Comprehensive conservation planning to protect biodiversity and ecosystem services in Canadian boreal regions under a warming climate and increasing

exploitation. *Biol Conserv* 2010; **143**: 1571–1586.

Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA, Oakley BB, Parks DH, Robinson CJ, Sahl JW, Stres B, Thallinger GG, Van Horn DJ, Weber CF. Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl Environ Microbiol* 2009; **75**: 7537–7541.

Schmidt SK, Costello EK, Nemergut DR, Cleveland CC, Reed SC, Weintraub MN, Meyer AF, Martin AM. Biogeochemical consequences of rapid microbial turnover and seasonal succession in soil. *Ecology* 2007; **88**: 1379–1385

Schütte UME, Abdo Z, Foster J, Ravel J, Bunge J, Solheim B, Forney LJ. Bacterial diversity in a glacier foreland of the high Arctic. *Mol Ecol* 2010; **19**: 54–66.

Shade A, Read JS, Youngblut ND, Fierer N, Knight R, Kratz TK, Lottig NR, Roden EE, Stanley EH, Stombaugh J, Whitaker RJ, Wu CH, McMahon KD. Lake microbial communities are resilient after a whole-ecosystem disturbance. *ISME J* 2012; **6**: 2153–2167.

Shade A, Jones SE, Gregory Caporaso J, Handelsman J, Knight R, Fierer N, Gilbert JA. Conditionally rare taxa disproportionately contribute to temporal changes in microbial diversity. *MBio* 2014; **5**: e01371-14.

Shelley F, Grey J, Trimmer M. Widespread methanotrophic primary production in lowland chalk rivers. *Proc R Soc B Biol Sci* 2014; **281**: 20132854.

Shipley B. Cause and correlation in biology: a user's guide to path analysis, structural equations and causal inference with R. 2016; Cambridge University Press: Cambridge, UK.

Shulman MJ, Ogden JC, Ebersole JP, McFarland WN, Miller SL, Wolf NG. Priority effects in the recruitment of juvenile coral reef fishes. *Ecology* 1983; **64**: 1508–1513.

Shurin JB, Arnott SE, Hillebrand H, Longmuir A, Pinel-Alloul B, Winder M, Yan ND. Diversity-stability relationship varies with latitude in zooplankton. *Ecol Lett* 2007; **10**: 127–134.

Sinsabaugh RL, Moorhead DL. Resource allocation to extracellular enzyme production: A

model for nitrogen and phosphorus control of litter decomposition. *Soil Biol Biochem* 1994; **26**: 1305–1311.

Sinsabaugh RL. Phenol oxidase, peroxidase and organic matter dynamics of soil. *Soil Biol Biochem* 2010; **42**: 391–404.

Sinsabaugh RL, Findlay S, Franchini P, Fischer D. Enzymatic analysis of riverine bacterioplankton production. *Limnol Oceanogr* 1997; **42**: 29–38.

Smith AP, Marín-Spiotta E, Balser T. Successional and seasonal variations in soil and litter microbial community structure and function during tropical postagricultural forest regeneration: A multiyear study. *Glob Chang Biol* 2015; **21**: 3532–3547.

Sobek S, Tranvik LJ, Prairie YT, Kortelainen P, Cole JJ. Patterns and regulation of dissolved organic carbon: An analysis of 7,500 widely distributed lakes. *Limnol Oceanogr* 2007; **52**: 1208–1219.

Solan M, Batty P, Bulling MT, Godbold JA. How biodiversity affects ecosystem processes: Implications for ecological revolutions and benthic ecosystem function. *Aquat Biol* 2008; **2**: 289–301.

Solieri L, Dakal TC, Giudici P. Next-generation sequencing and its potential impact on food microbial genomics. *Ann Microbiol* 2013; **63**: 21–37.

Sollins P, Homann P, Caldwell BA. Stabilization and destabilization of soil organic matter: Mechanisms and controls. *Geoderma* 1996; **74**: 65–105.

Solomon CT, Carpenter SR, Clayton MK, Cole JJ, Coloso JJ, Pace ML, Vander Zanden MJ, Weidel BC. Terrestrial, benthic, and pelagic resource use in lakes: Results from a three-isotope Bayesian mixing model. *Ecology* 2011; **92**: 1115–1125.

Solomon CT, Jones SE, Weidel BC, Buffam I, Fork ML, Karlsson J, Larsen S, Lennon JT, Read JS, Sadro S, Saros JE. Ecosystem Consequences of Changing Inputs of Terrestrial Dissolved Organic Matter to Lakes: Current Knowledge and Future Challenges. *Ecosystems* 2015; **18**: 376–389.

Spencer RGM, Bolton L, Baker A. Freeze/thaw and pH effects on freshwater dissolved organic matter fluorescence and absorbance properties from a number of UK locations. *Water*

Res 2007; **41**: 2941–2950.

Stachowicz JJ, Bruno JF, Duffy JE. Understanding the Effects of Marine Biodiversity on Communities and Ecosystems. *Annu Rev Ecol Evol Syst* 2007; **38**: 739–766.

Stelzer RS, Heffernan J, Likens GE. The influence of dissolved nutrients and particulate organic matter quality on microbial respiration and biomass in a forest stream. *Freshw Biol* 2003; **48**: 1925–1937.

Suberkropp K. Relationships between growth and sporulation of aquatic hyphomycetes on decomposing leaf litter. *Mycol Res* 1991; **95**: 843–850.

Swan BK, Ehrhardt CJ, Reifel KM, Moreno LI, Valentine DL. Archaeal and bacterial communities respond differently to environmental gradients in anoxic sediments of a california hypersaline lake, the Salton Sea. *Appl Environ Microbiol* 2010; **76**: 757–768.

Szkokan-Emilson EJ, Wesolek B., Gunn JM. Terrestrial organic matter as subsidies that aid in the recovery of macroinvertebrates in industrially damaged lakes. *Ecol Appl* 2011; **21**: 2082–2093.

Tanentzap AJ, Szkokan-Emilson EJ, Kielstra BW, Arts MT, Yan ND, Gunn JM. Forests fuel fish growth in freshwater deltas. *Nat Commun* 2014; **5**: 4077.

Tanentzap AJ, Szkokan-Emilson EJ, Desjardins CM, Orland C, Yakimovich K, Dirszowsky R, Mykytczuk N, Basiliko N, Gunn J. Bridging between litterbags and whole-ecosystem experiments: A new approach for studying lake sediments. *J Limnol* 2017; **76**: 431–437.

Tank JL, Dodds WK. Nutrient limitation of epilithic and epixylic biofilms in ten North American streams. *Freshw Biol* 2003; **48**: 1031–1049.

Tardy V, Mathieu O, Lévêque J, Terrat S, Chabbi A, Lemanceau P, Ranjard L, Maron PA. Stability of soil microbial structure and activity depends on microbial diversity. *Environ Microbiol Rep* 2014; **6**: 173–183.

Taylor BR, Parkinson D, Parsons WFJ. Nitrogen and lignin content as predictors of litter decay rates: a microcosm test. *Ecology* 1989; **70**: 97–104.

Tedersoo L, Bahram M, Põlme S, Kõljalg U, Yorou NS, Wijesundera R, Ruiz LV, Vasco-Palacios AM, Thu PQ, Suija A, Smith ME, Sharp C, Saluveer E, Saitta A, Rosas M, Riit

T, Ratkowsky D, Pritsch K, Põldmaa K, Piepenbring M, Phosri C, Peterson M, Parts K, Pärtel K, Otsing E, Nouhra E, Njouonkou AL, Nilsson RH, Morgado LN, Mayor J, May TW, Majuakim L, Lodge DJ, Lee S, Larsson KH, Kohout P, Hosaka K, Hiiesalu I, Henkel TW, Harend H, Guo LD, Greslebin A, Grelet G, Geml J, Gates G, Dunstan W, Dunk C, Drenkhan R, Dearnaley J, De Kesel A, Dang T, Chen X, Buegger F, Brearley FQ, Bonito G, Anslan S, Abell S, Abarenkov K. Global diversity and geography of soil fungi. *Science* 2014; **346**: 1256688.

Tessler M, Neumann JS, Afshinnekoo E, Pineda M, Hersch R, Velho LFM, Segovia BT, Lansac-Toha FA, Lemke M, DeSalle R, Mason CE, Brugler M. Large-scale differences in microbial biodiversity discovery between 16S amplicon and shotgun sequencing. *Sci Rep* 2017; **7**: 6589.

Thébault E, Loreau M. Trophic interactions and the relationship between species diversity and ecosystem stability. *Am Nat* 2005; **166**: 95–114.

Thode HG. Sulphur isotopes in nature and the environment: An Overview. 1991; In Krouse HR, Grinenko VA (Eds.), *Stable Isotopes: Natural and Anthropogenic Sulfur in the Environment*. Wiley; pp. 1–26.

Thomas MK, Fontana S, Reyes M, Kehoe M, Pomati F. The predictability of a lake phytoplankton community, over time-scales of hours to years. *Ecol Lett* 2018; **21**: 619–628

Thompson LR, Sanders JG, McDonald D, Amir A, Ladau J, Locey KJ, Prill RJ, Tripathi A, Gibbons SM, Ackermann G, Navas-Molina JA, Janssen S, Kopylova E, Vázquez-Baeza Y, González A, Morton JT, Mirarab S, Zech Xu Z, Jiang L, Haroon MF, Kanbar J, Zhu Q, Jin Song S, Kosciolek T, Bokulich NA, Lefler J, Brislawn CJ, Humphrey G, Owens SM, Hampton-Marcell J, Berg-Lyons D, McKenzie V, Fierer N, Fuhrman JA, Clauset A, Stevens RL, Shade A, Pollard KS, Goodwin KD, Jansson JK, Gilbert JA, Knight R, Earth Microbiome Project Consortium. A communal catalogue reveals Earth's multiscale microbial diversity. *Nature* 2017; **551**: 457–463.

Tilman D. Resource competition and community structure. 1982; Princeton University Press:

Princeton, NJ.

Tilman D. The ecological consequences of changes in biodiversity: a search for general principles. *Ecology* 1999; **80**: 1455–1474.

Tilman D, Lehman CL, Bristow CE. Diversity-Stability Relationships: Statistical Inevitability or Ecological Consequence? *Am Nat* 1998; **151**: 277–282.

Tilman D, Reich PB, Isbell F. Biodiversity impacts ecosystem productivity as much as resources, disturbance, or herbivory. *Proc Natl Acad Sci* 2012; **109**: 10394–10397.

Tilman D, Isbell F, Cowles JM. Biodiversity and Ecosystem Functioning. *Annu Rev Ecol Evol Syst* 2014; **45**: 471–493.

Tranvik LJ, Downing JA, Cotner JB, Loiselle SA, Striegl RG, Ballatore TJ, Dillon P, Finlay K, Fortino K, Knoll LB, Kortelainen PL, Kutser T, Larsen S, Laurion I, Leech DM, McCallister SL, McKnight DM, Melack JM, Overholt E, Porter JA, Prairie Y, Renwick WH, Roland F, Sherman BS, Schindler DW, Sobek S, Tremblay A, Vanni MJ, Verschoor AM, von Wachenfeldt E, Weyhenmeyer GA. Lakes and reservoirs as regulators of carbon cycling and climate. *Limnol Oceanogr* 2009; **54**: 2298–2314.

Tripathi BM, Stegen JC, Kim M, Dong K, Adams JM, Lee YK. Soil pH mediates the balance between stochastic and deterministic assembly of bacteria. *ISME J* 2018; **12**: 1072–1083.

Trivedi P, Delgado-Baquerizo M, Trivedi C, Hu H, Anderson IC, Jeffries TC, Zhou J, Singh BK. Microbial regulation of the soil carbon cycle: Evidence from gene-enzyme relationships. *ISME J* 2016; **10**: 2593–2604.

Tylianakis JM, Laliberté E, Nielsen A, Bascompte J. Conservation of species interaction networks. *Biol Conserv* 2010; **143**: 2270–2279.

Vass M, Langenheder S. The legacy of the past: Effects of historical processes on microbial metacommunities. *Aquat Microb Ecol* 2017; **79**: 13–19.

Vellend M. Conceptual Synthesis in Community Ecology. *Q Rev Biol* 2010; **85**: 183–206.

Venail PA, Vives MJ. Positive Effects of Bacterial Diversity on Ecosystem Functioning Driven by Complementarity Effects in a Bioremediation Context. *PLoS One* 2013; **8**: e72561.

Venter JC, Remington K, Heidelberg JF, Halpern AL, Rusch D, Eisen JA, Wu D, Paulsen I, Nelson KE, Nelson W, Fouts DE, Levy S, Knap AH, Lomas MW, Nealson K, White O, Peterson J, Hoffman J, Parsons R, Baden-Tillson H, Pfannkoch C, Rogers YH, Smith HO. Environmental genome shotgun sequencing of the Sargasso Sea. *Science* 2004; **304**: 66–74.

Vogt RJ, Romanuk TN, Kolasa J. Species richness-variability relationships in multi-trophic aquatic microcosms. *Oikos* 2006; **113**: 55–66.

Vollenweider RA, Kerekes J. Eutrophication of Waters. Monitoring, Assessment and Control. 1982; Organization for Economic Co-Operation and Development (OECD), Paris.

Wagg C, Dudenhöffer JH, Widmer F, van der Heijden MGA. Linking diversity, synchrony and stability in soil microbial communities. *Funct Ecol* 2018; **32**: 1280–1292.

Wang B, Lerdau M, He Y. Widespread production of nonmicrobial greenhouse gases in soils. *Glob Chang Biol* 2017; **23**: 4472–4482.

Ward CP, Cory RM. Complete and Partial Photo-oxidation of Dissolved Organic Matter Draining Permafrost Soils. *Environ Sci Technol* 2016; **50**: 3545–3553.

Weishaar JL, Aiken GR, Bergamaschi BA, Fram MS, Fujii R, Mopper K. Evaluation of specific ultraviolet absorbance as an indicator of the chemical composition and reactivity of dissolved organic carbon. *Environ Sci Technol* 2003; **37**: 4702–4708.

Weiss S, Xu ZZ, Peddada S, Amir A, Bittinger K, Gonzalez A, Lozupone C, Zaneveld JR, Vázquez-Baeza Y, Birmingham A, Hyde ER, Knight R. Normalization and microbial differential abundance strategies depend upon data characteristics. *Microbiome* 2017; **5**: 27.

Wertz S, Degrange V, Prosser JI, Poly F, Commeaux C, Guillaumaud N, Le Roux X. Decline of soil microbial diversity does not influence the resistance and resilience of key soil microbial functional groups following a model disturbance. *Environ Microbiol* 2007; **9**: 2211–2219.

Wetzel RG. Sources and regulatory functions of dissolved organic matter in freshwater ecosystems. *Hydrobiologia* 1992; **229**: 181–198.

Wetzel RG. Death, detritus and energy flow in aquatic ecosystems. *Freshw Biol* 1995; **33**: 83–89.

Wetzel R. Limnology: Lake and River Ecosystems. 2001; Academic Press: New York, NY.

White TJ, Bruns T, Lee SJ, Taylor JL. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. *PCR protocols: a guide to methods and applications* 1990; **18**: 315-22.

Wieder WR, Grandy AS, Kallenbach CM, Bonan GB. Integrating microbial physiology and physio-chemical principles in soils with the MIcrobial-MIneral Carbon Stabilization (MIMICS) model. *Biogeosciences* 2014; **11**: 3899–3917.

Wiens JJ, Donoghue MJ. Historical biogeography, ecology and species richness. *Trends Ecol Evol* 2004; **19**: 639–644.

Williams CJ, Yamashita Y, Wilson HF, Jaffe R, Xenopoulos MA. Unraveling the role of land use and microbial activity in shaping dissolved organic matter characteristics in stream ecosystems. *Limnol Oceanogr* 2010; **55**: 1159–1171.

Williams RJ, Howe A, Hofmockel KS. Demonstrating microbial co-occurrence pattern analyses within and between ecosystems. *Front Microbiol* 2014; **5**: 358.

Williamson CE, Morris DP, Pace ML, Olson OG. Dissolved organic carbon and nutrients as regulators of lake ecosystems: Resurrection of a more integrated paradigm. *Limnol Oceanogr* 1999; **44**: 795–803.

Woodward FI, Diament AD. Functional approaches to predicting the ecological effects of global change. *Funct Ecol* 1991; **5**: 202–212.

Wooley JC, Godzik A, Friedberg I. A primer on metagenomics. *PLoS Comput Biol* 2010; **6**: e1000667.

Wurtsbaugh, W, Cole J, McKnight D, Brezonik P, MacIntyre S, Potter K. Report of The Workshop on Emerging Research Issues for Limnology: The study of Inland Waters. 2002; American Society of Limnological Oceanography: Boulder, CO.

Wurzbacher CM, Bärlocher F, Grossart HP. Fungi in lake ecosystems. *Aquat Microb Ecol* 2010; **59**: 125–149.

Yachi S, Loreau M. Biodiversity and ecosystem productivity in a fluctuating environment: The insurance hypothesis. *Proc Natl Acad Sci* 1999; **96**: 1463–1468.

Yakimovich KM, Szkokan-Emilson EJ, Carson MA, Tanentzap AJ, Basiliko N, Mykytczuk NCS. Plant litter type dictates microbial communities responsible for greenhouse gas production in amended lake sediments. *Front Microbiol* 2018; **9**: 2662.

Yu Y, Lee C, Kim J, Hwang S. Group-specific primer and probe sets to detect methanogenic communities using quantitative real-time polymerase chain reaction. *Biotechnol Bioeng* 2005; **89**: 670–679.

Zeglin LH. Stream microbial diversity in response to environmental changes: Review and synthesis of existing research. *Front Microbiol* 2015; **6**: 454.

Zeileis A, Hothorn T. Diagnostic checking in regression relationships. *R News* 2002; **2**: 7–10.

Zelezniak A, Andrejev S, Ponomarova O, Mende DR, Bork P, Patil KR. Metabolic dependencies drive species co-occurrence in diverse microbial communities. *Proc Natl Acad Sci* 2015; **112**: 6449–6454.

Zhang Y, Zhang X, Liu X, Xiao Y, Qu L, Wu L, Zhou J. Microarray-based analysis of changes in diversity of microbial genes involved in organic carbon decomposition following land use/cover changes. *FEMS Microbiol Lett* 2007; **266**: 144–151.

Zinger L, Boetius A, Ramette A. Bacterial taxa-area and distance-decay relationships in marine environments. *Mol Ecol* 2014; **23**: 954–964.

Zumsteg A, Luster J, Göransson H, Smittenberg RH, Brunner I, Bernasconi SM, Zeyer J, Frey B. Bacterial, Archaeal and Fungal Succession in the Forefield of a Receding Glacier. *Microb Ecol* 2012; **63**: 552–564.

Zuppinger-Dingley D, Schmid B, Petermann JS, Yadav V, De Deyn GB, Flynn DFB. Selection for niche differentiation in plant. *Nature* 2014; **515**: 108–111.

Zwart JA, Craig N, Kelly PT, Sebestyen SD, Solomon CT, Weidel BC, Jones SE. Metabolic and physiochemical responses to a whole-lake experimental increase in dissolved organic carbon in a north-temperate lake. *Limnol Oceanogr* 2016; **61**: 723–734.