Université de Montréal

Eco-evolutionary dynamics of microbial communities in disturbed freshwater ecosystems

par Naíla Barbosa da Costa

Département de sciences biologiques, Faculté des Arts et Sciences

Thèse présentée
en vue de l'obtention du grade de Philosophiae Doctor (Ph.D.)
en sciences biologiques

Août 2022

Université de Montréal
Département de sciences biologiques, Faculté des Arts et Sciences

Cette thèse intitulée

Eco-evolutionary dynamics of microbial communities in disturbed freshwater ecosystems

Présentée par

Naíla Barbosa da Costa

A été évaluée par un jury composé des personnes suivantes

Dre. Sandra Binning
Présidente-rapporteuse

Dr. Jesse Shapiro
Directeur de recherche

Dr. Gregor Fussmann
Co-directeur

Dr. Steven Kembell
Membre du jury

Dre. Laura Hug
Examinatrice externe

Résumé

L'intensification de l'activité agricole depuis la deuxième moitié du 20e siècle, notamment l'utilisation de produits agrochimiques dans les bassins versants, a affecté la qualité des ressources d'eau douce. Des traces de produits agrochimiques, tels que les pesticides et les engrais, sont transportées par ruissellement de surface ou lixiviation, provoquant des effets directs ou indirects sur les organismes aquatiques. Se trouvant à la base des réseaux trophiques aquatiques, les micro-organismes sont des habitants indispensables dans les écosystèmes d'eau douce, où ils jouent également un rôle important pour les services écosystémiques en tant que propulseurs des cycles biogéochimiques. En faisant partie de l'écosystème, les communautés bactériennes sont susceptibles aux perturbations anthropiques croissantes qui se déroulent dans leurs milieux. Le but principal de cette thèse est d'étudier l'effet de perturbations agricoles simulées sur les bactéries d'eau douce par une approche expérimentale avec des réservoirs extérieurs (mésocosmes) et en utilisant le séquençage d'ADN à haut débit.

Des mésocosmes ont été remplis de 1 000 litres d'eau provenant d'un lac bien préservé et, ensuite, ont été traités avec des pesticides largement utilisés au monde en combinaison avec des engrais. Les trois études présentées dans cette thèse explorent les réponses du bactérioplancton dans cette même expérience sous différents angles : la première (chapitre II) s'est concentrée sur les réponses écologiques des communautés bactériennes à de différentes combinaisons de produits agrochimiques; la deuxième (chapitre III) a examiné si les gènes de résistance aux antibiotiques pourraient changer le succès d'espèces soumises à une grave contamination par un herbicide et, finalement, la troisième (chapitre IV) a suivi les altérations évolutives parmi les espèces ayant des réponses écologiques similaires par rapport au traitement avec l'herbicide.

En mettant l'accent sur la réaction des communautés exposées à un mélange de produits agrochimiques, le chapitre II complémente des études écotoxicologiques, qui se concentrent traditionnellement sur les réponses d'une seule espèce à des produits chimiques isolés. Les mésocosmes ont été exposés à de différentes concentrations d'un herbicide à base de glyphosate et d'un insecticide néonicotinoïde, séparés ou en combinaison, en plus d'apports faibles ou élevés en nutriments. Le séquençage des amplicons du gène de l'ARNr 16S et la prédiction des variantes de séquences ont étés

faits pour étudier la diversité taxonomique, ainsi que le profilage de l'utilisation microbienne des sources de carbone pour décrire les changements de diversité fonctionnelle à travers le temps. Les résultats ont révélé que la stabilité des communautés microbiennes varie en fonction du type et de l'intensité de la perturbation. Bien que les communautés bactériennes n'aient pas réagi à l'introduction de l'insecticide ou d'engrais, elles sont modifiées de manière intensive sous des concentrations élevées de l'herbicide à base de glyphosate. Des aspects distincts de la diversité des communautés ont réagi différemment aux perturbations : alors que la composition fonctionnelle est restée stable face aux perturbations, la composition taxonomique au niveau taxonomique le plus fin a été sensible au glyphosate et résiliente aux échelles taxonomiques plus larges (c'est-à-dire, du genre au phylum). Ces résultats soulignent la complexité des réponses écologiques et fournissent des évidences de la redondance fonctionnelle concernant l'utilisation des sources de carbone dans les communautés microbiennes.

Le chapitre III a testé l'hypothèse selon laquelle les gènes de résistance aux antibiotiques, en particulier les pompes d'efflux, favorisent la survie des bactéries en présence de l'herbicide à base de glyphosate. Cette hypothèse n'a été confirmée que par des études expérimentales en laboratoire avec des cultures bactériennes et plus récemment dans les microbiomes du sol. C'était donc la première fois que cette hypothèse a été testée dans un système aquatique. Au chapitre II, on a observé que l'herbicide à base de glyphosate favorisait la domination de nombreux taxons de l'embranchement des protéobactéries, dont *Agrobacterium*, un genre qui code pour l'enzyme cible du glyphosate appartenant à la classe des résistants. Cependant, d'autres espèces codant pour la classe de l'enzyme sensible au glyphosate étaient également favorisées, ce qui implique le rôle d'autres mécanismes de résistance. Dans le chapitre III, les analyses de métagénomes et des génomes assemblés par métagénomes ont révélé une augmentation de la fréquence de gènes de résistance aux antibiotiques après l'administration de fortes doses de l'herbicide. D'ailleurs, l'abondance relative des espèces présentes après qu'une forte dose de l'herbicide a été administrée était mieux prédite par la présence de gènes d'efflux d'antibiotiques dans leur génome que par la présence du gène codant pour l'enzyme résistante au glyphosate. Ces résultats

renforcent les études récentes et contribuent aux premières évidences provenant des communautés bactériennes d'eau douce.

L'objectif du chapitre IV était de vérifier si les bactéries ayant la même réponse écologique à la contamination par l'herbicide à base de glyphosate présenteraient également des réponses évolutives similaires. En plus, ce chapitre avait pour but de contribuer aux preuves expérimentales du modèle de l'écotype stable, un modèle proéminent sur l'évolution et l'origine de la diversité dans les espèces bactériennes. On a supposé que les espèces favorisées par l'herbicide subiraient des balayages sélectifs éliminant la variation génétique dans le génome, comme le prédit le modèle évolutif de l'écotype stable. Pour tester cette hypothèse, des polymorphismes nucléotidiques ont été quantifiés au sein des populations bactériennes au cours du temps dans 12 populations bien représentées dans le séquençage métagénomique qui a été fait dans le chapitre III. Contrairement à ce que l'on attendait, les populations écologiquement prospères ont montré une variété de réponses évolutives et la diversité n'a été supprimée que dans quelques-unes d'entre elles. Les résultats montrent que d'autres mécanismes évolutifs qui maintiennent la variation génétique, tels que des balayages sélectifs à l'échelle du gène plutôt qu'à l'échelle du génome, peuvent être plus souvent impliqués dans le succès des espèces qui survivent au stress anthropique.

Mis ensemble, ces résultats soulignent la complexité des réponses bactériennes face à une perturbation anthropique au niveau des communautés, des populations, des gènes et des allèles. Les connaissances apportées par cette thèse peuvent améliorer les évaluations des risques de déversements accidentels en eau douce. Le changement permanent à des niveaux taxonomiques fins et la sélection croisée pour les gènes de résistance aux antibiotiques en présence de concentrations élevées d'herbicides indiquent des risques qui devraient être mieux compris par rapport à leur prédominance et les mécanismes qui les causent. D'ailleurs, la dynamique évolutive décrite ici sur une échelle de temps de courte durée fournit des données pour soutenir une importante théorie sur la différenciation et la spéciation bactériennes.

Mots clés : microbiome d'eau douce, contamination agricole, écologie des communautés, biodiversité, stabilité, métagénomique, résistance aux antibiotiques, évolution rapide.

Abstract

Agriculture intensification in the second half of the 20th century, particularly the use of agrochemicals within watersheds, has affected freshwater quality. Traces of agrochemicals, such as pesticides and fertilizers, reach freshwater systems through runoff or leaching, causing direct or indirect effects on aquatic organisms. Microorganisms are essential inhabitants of aquatic systems as they are at the foundation of food webs and play roles in ecosystem functioning as important drivers of biogeochemical cycles. By being part of the ecosystem, bacterial communities are subject to the increasing anthropogenic perturbations in their environment. The main objective of this thesis is to investigate the effect of simulated agricultural perturbations on freshwater bacteria through an experimental approach with outdoor tanks (mesocosms) and using high-throughput DNA sequencing.

Mesocosms were filled with 1,000 L of water from a pristine freshwater lake and treated with widely used pesticides in combination with fertilizers. The three main studies in this thesis explored the bacterioplankton responses in this experiment through different angles: the first study (chapter II) focused on ecological responses to a combination of agrochemicals; the second (chapter III) explored how changes in antibiotic resistance genes could explain the ecological success of species facing severe herbicide contamination and the third study (chapter IV) tracked evolutionary changes among species with similar ecological responses to the herbicide treatment.

Chapter II aimed to complement ecotoxicological studies, that traditionally focus on single species responses to individual chemicals, by focusing on communities exposed to a mixture of agrochemicals, as typically observed in nature. For that, the mesocosms were exposed to different concentrations of a glyphosate-based herbicide and a neonicotinoid insecticide, isolated or in combination, in addition to low or high nutrient inputs. Sequencing of 16S rRNA gene amplicons and inference of amplicon sequence variants were done to study taxonomic diversity, as well as profiling microbial use of carbon sources to describe functional diversity changes through time. The results revealed that the stability of microbial communities varies according to the type and intensity of the disturbance. The highest dose of the glyphosate-based herbicide was the major driver of ecological responses within bacterial communities, which were not altered

by the insecticide nor by nutrient fertilization. Distinct aspects of community diversity responded differently to perturbation: while functional composition remained stable in face of disturbances, taxonomic composition was sensitive to glyphosate at the finest taxonomic level and resilient at higher taxonomic units (i.e. genus to phylum). These results highlight the complexity of ecological responses and provide evidence of functional redundancy regarding the use of carbon sources in these communities.

Chapter III tested the hypothesis that antibiotic resistance genes, particularly efflux pumps, would favour bacterial survival in the presence of the glyphosate-based herbicide. This hypothesis has only been confirmed through experimental laboratory studies with bacterial cultures and more recently in soil microbiomes, it was thus the first time it was tested in an aquatic system. As observed in chapter II, glyphosate-based herbicide favoured the dominance of many taxa of the phylum Proteobacteria, including *Agrobacterium*, a genus that encodes the glyphosate-resistant target enzyme. However, other species encoding the glyphosate-sensitive version of the enzyme were also favoured, implying other resistance mechanisms. In chapter III, the analysis of metagenomes and metagenome-assembled genomes revealed an increased frequency of antibiotic resistance genes following high doses of the herbicide. Additionally, the relative abundance of species after a severe herbicide pulse was better predicted by the presence of antibiotic efflux genes in their genome than by the presence of the gene encoding the resistant glyphosate target enzyme. These results reinforce recent studies and contribute to the first evidence from freshwater bacterial communities.

The goal of chapter IV was to test if bacteria with the same ecological response to the contamination with the glyphosate-based herbicide would also show similar evolutionary responses. Furthermore, this chapter aimed to contribute to experimental evidence to the stable ecotype model, a prominent model on the evolution and origin of diversity in bacterial species. If assumptions of the stable ecotype model were confirmed by the experiment, species favoured by the herbicide would experience selective sweeps purging genetic variation across the genome. To test this hypothesis, single nucleotide variants were quantified within bacterial populations over time in 12 populations well-represented in the metagenomic sequencing that was performed in chapter III. Differently than expected, ecologically successful populations showed a variety of evolutionary

responses and diversity was purged only in a few of them. The results show that other evolutionary mechanisms that maintain genetic variation, such as gene-wide specific sweeps rather than genome-wide sweeps, may be more often involved in the success of species surviving anthropogenic stress.

Together, these results highlight the complexity of bacterial responses in the face of an anthropogenic disturbance at the level of communities, populations, genes, and alleles. The knowledge provided by this thesis may improve assessments of the potential risks of accidental spills in freshwater. The permanent change at fine taxonomic levels and the cross-selection for antibiotic resistance genes in the presence of high concentrations of herbicide indicate risks that should be better understood regarding their predominance and causing mechanisms. Moreover, the evolutionary dynamics here described in a short-term time scale provide observational data to support a theoretical background on bacterial differentiation and speciation.

Keywords: freshwater microbiome, agricultural contamination, community ecology, biodiversity, stability, metagenomics, antibiotic resistance, rapid evolution

Table of contents

List of figures

Main figures

CHAPTER I

CHAPTER II

CHAPTER III

CHAPTER IV

List of tables

Main tables

CHAPTER III

Supplementary tables

CHAPTER II

List of acronyms and abbreviations

ARG: antibiotic resistance gene

ABC: ATP-binding cassette

ANI: nucleotide identity

ARO: antibiotic resistance ontology

ASV: amplicon sequence variant

ATP: adenosine triphosphate

AWCD: average well color development

BLAST: basic local alignment search tool

bp: base pair

C : Celsius

CARD: Comprehensive Antibiotic Resistance Database

CCME: Canadian Council of Ministers of the Environment

cm: centimeter

CREATE: Collaborative Research and Training Experience program

CRISPR: clustered regularly interspaced short palindromic repeats

CV: coefficient of variation

df: degrees of freedom

DNA: deoxyribonucleic acid

EAA: estimated absolute abundance

EDF: effective degrees of freedom

EDTA: ethylenediaminetetraacetic acid

EFSA: European Food Safety Agency

EPA: Environmental Protection Agency

EPSPS: enolpyruvylshikimate-3-phosphate synthase

FAO: Food and Agriculture Organization of the United Nations

g: gram

GAM: generalized additive model

GAMM: generalized additive mixed model

GBH: glyphosate-based herbicide

GRIL: *Groupe de Recherche Interuniversitaire en Limnologie*

JSD: Jensen-Shannon divergence

K: potassium

L: litter

LEAP: Large Experimental Array of Ponds

log: logarithm

m: meters

Mbp: megabase pair

MAF: minor allele frequency

MAG: metagenome-assembled genome

max: maximum

maxEE: maximum expected error

mg: microgram

min: minute

mM: milimolar

N: nitrogen

NCBI: National Center for Biotechnology Information

NSERC: Natural Sciences and Engineering Research Council

O: oxygen

OTU: operational taxonomic unit

P: phosphorus

PCoA: principal coordinates analysis

PERMANOVA: permutational multivariate ANOVA

PERMDISP: Permutational analyses of multivariate dispersions

pH: potential of hydrogen

POEA: polyethoxylenamine

PRC: polymerase chain reaction

PRC: principal response curve

pRDA: partial redundancy analysis

QCBS: Québec Centre for Biodiversity Science

RAC: regulatory acceptable concentration

RDA: redundancy analysis

RGI: resistance gene identifier

rRNA: ribosomal ribonucleic acid

SE: standard error

SNV: single nucleotide variant

TN: total nitrogen

TP: total phosphorus

USA: United States of America

À minha mãe e aos meus avós,
por tudo que me foi compartilhado
verticalmente e horizontalmente

Acknowledgments

Paraphrasing the proverb, one could say it takes a village to finish a PhD, and if the student comes from abroad like me, it takes many villages sometimes separated by continental distances. I am grateful to have had the support of an incredible international community in Montreal, of my family and friends in Brazil, and also of those who welcomed me into their home, sharing their culture, rights and claims.

I acknowledge the first inhabitants of these lands, the more than 50 peoples in Canada and about 200 peoples in Brazil, that never gave up protecting their spaces and that preserved the two countries that I call home so impressively rich and strong. I learned to speak the languages of the colonizers, but the truth about our countries was first told in many other native languages whose conservation will also preserve our history.

I am thankful to my supervisor, who has been for the last six years the most supportive, understanding and encouraging supervisor I could ever imagine. I am also thankful to my co-supervisor, my collaborators and my co-workers, particularly those who became my friends during this journey.

I am grateful for the financial support provided by the NSERC-CREATE program through GRIL, without which I would not have been able to move to Canada, and for the support additionally provided through FRQNT. I am also grateful to the GRIL and QCBS communities that helped me to integrate into a new environment.

Finally, I am extremely grateful to have had the support of my partner and my mother since I took the decision to start a PhD until the last word I wrote in this thesis. I am lucky to be able to count on people like you on my daily life. I became a different person and a better scientist along my PhD and I owe that to all those mentioned here.

CHAPTER I : GENERAL INTRODUCTION

Agriculture in the Anthropocene and freshwater contamination

Land use intensification for human activities such as agriculture and urbanization have transformed natural landscapes across most of the Earth's ecosystems (Ellis et al., 2013; Foley et al., 2005). As a consequence, the planet currently faces major changes in climate (IPCC, 2022), biogeochemical cycles (Tranvik et al., 2009), as well as habitat loss and direct and indirect effects on biodiversity (Parmesan & Yohe, 2003), including environmental microbiomes (Kraemer et al., 2020; Zhu & Penuelas, 2020). The intensity and long-lasting consequences of human-driven global change led to the proposition of a new geologic epoch, the Anthropocene, as the anthropogenic impact on the planet in the last centuries is comparable in magnitude and significance to events that define transitions between other epochs in the Earth's geological history (Steffen, Crutzen, & McNeill, 2007; Zalasiewicz, Williams, Haywood, & Ellis, 2011).

Agricultural production represents the largest use of land on the planet, globally estimated to occupy about 40% of land surface with croplands and pastures (Campbell et al., 2017; Foley et al., 2005). Although the first evidence of agriculture dates about 11,000 ago (Lewis & Maslin, 2015) (Figure I.1), over the last century its intensification led to extensive environmental damages, such as the degradation of water quality, deforestation and greenhouse gas emissions (Campbell et al., 2017; FAO, 2017; Pimentel & Pimentel, 1990). The efforts to increase agricultural productivity through technological advances, also known as the green revolution (Figure I.1), was responsible for tripling agricultural production from 1961 to 2011 worldwide. However, these gains were also accompanied by negative effects on the environment mainly due to intensified use of land, irrigation water and agrochemicals (FAO, 2017). Although the green revolution is usually justified by the world's growing food production needs (Gurdev S. Khush, 2001), practical solutions to achieve global food security and environmental sustainability have been proposed to address both challenges (Foley et al., 2011) and opposition to adopt them usually lean to economic and sociopolitical issues rather than their feasibility (Tollefson, 2019).

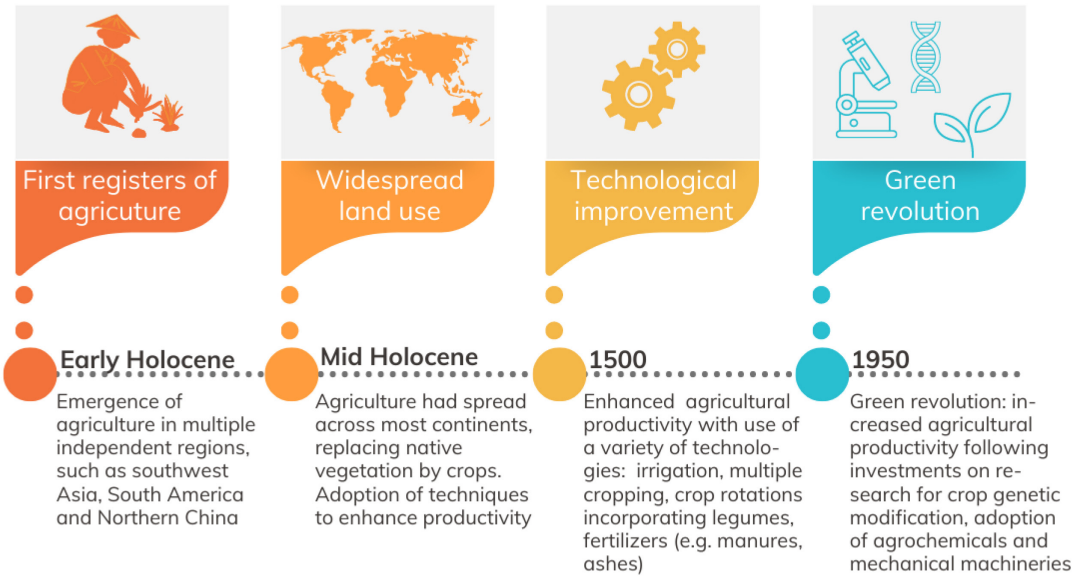


Figure I.1 Important landmarks in the history of agriculture (Ellis et al., 2013). Early Holocene started about 11,000 years ago and the mid-Holocene about 6,000 or 5,000 years ago. Illustration made with ©Canva (canva.com).

Mazor et al. (2018) identified a misalignment between global policy targets and research priorities on drivers of biodiversity loss, emphasizing the need to address the biodiversity impact of pollution in ecological and conservation studies. Pollution of different environments remains understudied relative to its projected high impact on ecosystem health (Mazor et al., 2018). For example, pollution with organic chemicals, which include agrochemicals such as pesticides, is a continental-scale problem threatening freshwater biodiversity in almost half of 4,000 European monitoring sites, whose chemical risks are strongly associated with the presence of upstream agriculture and urban areas (Malaj et al., 2014).

The intense use of agrochemicals, such as pesticides and fertilizers, raises concerns about human water security and freshwater biodiversity (Moss, 2008; Vörösmarty et al., 2010), as they can be mobilized from their point of application and transported to freshwater systems through leaching and surface runoff (M.-P. Hébert, Fugère, & Gonzalez, 2018; Wittmer et al., 2010) (Figure I.2). The use of pesticides, designed to control the occurrence of unwanted organisms, has increased in the last

decades, particularly since the adoption of genetically modified crops resistant to herbicides in the mid-1990s, which facilitated weeds control without harming crops and, in response to intensified herbicide use, resistant weeds started to be observed creating the need to each time increase herbicide application doses to their extermination (Mortensen, Egan, Maxwell, Ryan, & Smith, 2012). In 2019, the global annual application of herbicides surpassed 4 million tons (FAO, 2021b). Global annual application of fertilizers, which provide nutrients for faster growth and better yields of crops, also exceeds millions of tons (FAO, 2021a). Although pesticides and fertilizers are concomitantly found in surface waters (Altenburger, Backhaus, Boedeker, Faust, & Scholze, 2013), their interactive impact is rarely assessed and deserves more attention as organisms are often exposed to mixtures of contaminants rather than individual chemicals and their effects may be different when analyzed separately (Geyer, Smith, & Rettig, 2016; Relyea, 2009).

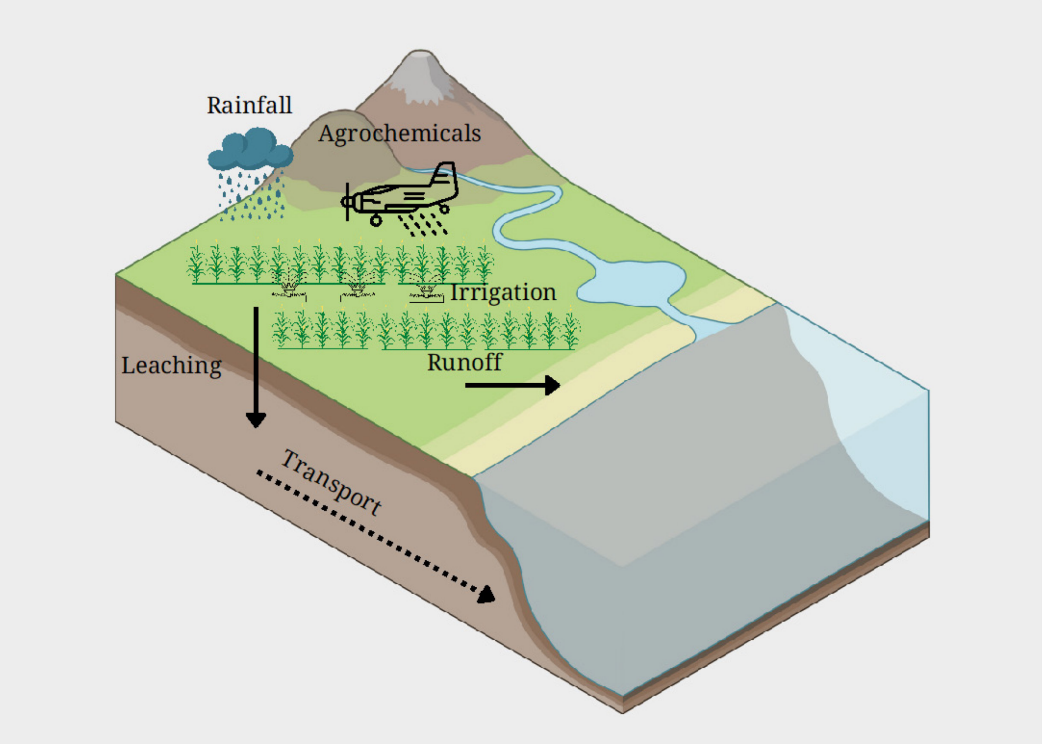


Figure I.2 Simple schematic of how agrochemicals reach freshwater resources, made with ©Canva (canva.com) and ©BioRender (biorender.com).

Freshwater microbes and agricultural pollution

Falkowski et al. (2008) claimed that microorganisms are the "engines that drive Earth's biogeochemical cycles" as they evolved protein complexes able to catalyze highly energetic redox reactions essential for life on Earth. These reactions are fundamental to the biogeochemical fluxes of chemical elements that compose essential biomolecules, such as carbohydrates, proteins, and nucleic acid (Figure I.3). Fluxes of carbon, hydrogen, oxygen, nitrogen and sulfur are mostly driven by "biological machines" encoded in microbial genomes (Falkowski et al., 2008). Phosphorus is also an important element of biological macromolecules that accumulates in the sediments of aquatic systems and whose bioavailability is associated with the enzymatic activity of microorganisms (Williamson, Saros, Vincent, & Smol, 2009). Freshwater resources, particularly lentic systems such as lakes and reservoirs, are important sinks of these major elements during their transport across the landscape, acting as hotspots for biogeochemical cycles (Cheng & Basu, 2017; Cole et al., 2007), where microorganisms play essential roles catalyzing chemical reactions.

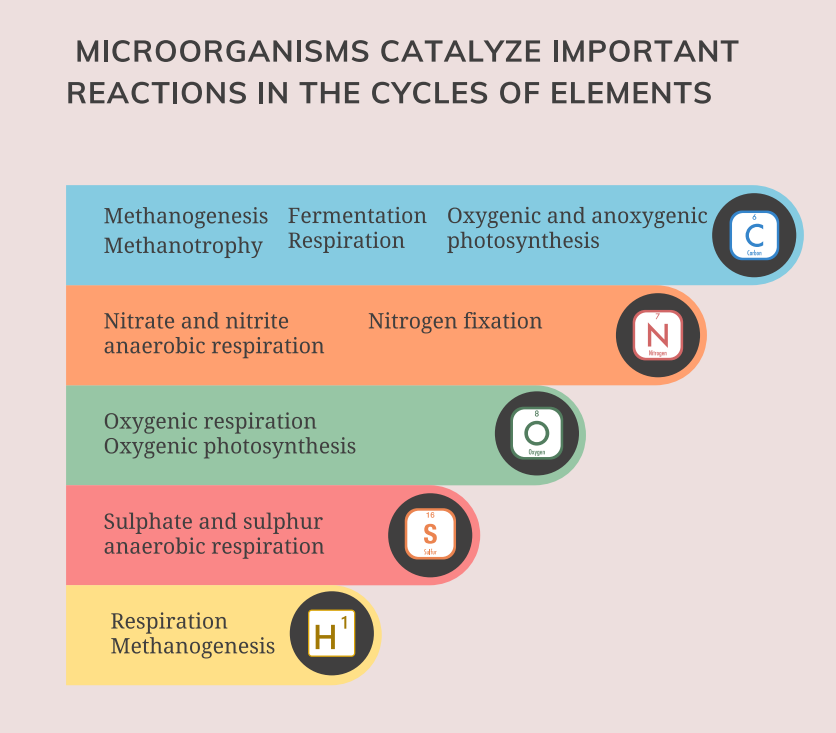


Figure I.3 Non-extensive list of reactions catalyzed by microorganisms organized by the element whose flux they participate (Falkowski et al., 2008). These elements, together

with phosphorus, are essential components of biomolecules. Illustration made with ©Canva (canva.com).

Freshwater resources represent a substantial percentage of the Canadian landscape, with almost 1 million lakes occupying about 10% of the country's territory and contributing about 20% of the world's freshwater stocks (Huot et al., 2019). Lakes provide significant ecosystem services to both humans and wildlife, such as oxygen and greenhouse gas production (Cole et al., 2007). However, as anthropogenic interferences increase worldwide, lake health and the ecosystem services they provide are susceptible to changes (Williamson, Dodds, Kratz, & Palmer, 2008), as are the microbial communities inhabiting lakes (Kraemer et al., 2020).

Human activity and especially intensive agriculture affects freshwater microbial communities with consequences for ecosystem processes. A common consequence of agricultural land use within watersheds is eutrophication, a process by which water bodies become enriched in nutrients and experience an excessive increase in primary production and loss of biodiversity (V. H. Smith & Schindler, 2009). Geographic regions with large areas occupied by surface waters are also more prone to agricultural development and thus to eutrophication (Pacheco, Roland, & Downing, 2014). Microbial community changes caused by eutrophication have pronounced ecosystem impacts with many cascading effects, such as the increased frequency and intensity of harmful cyanobacterial blooms (V. H. Smith & Schindler, 2009). Blooms are associated with decreased fish survival due to intoxication and hypoxia (i.e. the oxygen depletion in the hypolimnion, the deeper layer of stratified lakes) caused by increased bacterial respiration as the sediments become richer in organic matter (Scavia et al., 2014). Another important change of global concern is the enhanced emissions of methane, a potent greenhouse gas produced by methanogenic archaea and consumed by methanotrophic bacteria (Evans et al., 2019; Reis, Thottathil, & Prairie, 2022), whose emissions from lakes and impoundments are predicted to grow considerably in this century as eutrophication increases worldwide (Beaulieu, DelSontro, & Downing, 2019).

Bacterial communities generally respond rapidly to environmental changes, which can be explained by their high adaptability given their large mutation supply in large

population sizes (Shapiro & Polz, 2014). When faced with a disturbance, communities may remain stable (components of stability are explained in the next section) or oscillate in their species composition and/or function (Shade et al., 2012). In diverse microbial communities, some functions are usually redundant (Louca et al., 2016) and it is possible that, even after a change in species composition, functions remain stable. Nevertheless, the microbial ecology literature needs more empirical work defining microbial functional groups and their responses to disturbances, as well as their links with phylogeny. In other words, we lack research focusing on the relationship between microbial taxonomic groups and functional traits in communities facing environmental changes (Allison & Martiny, 2008).

Ecology and evolution in the response of microbes to environmental change

Both ecology and evolution play a role in microbial responses to environmental perturbations, and understanding their relative contribution is a fundamental question. Although evolutionary responses were traditionally considered slow compared to ecological responses, both can occur rapidly and simultaneously in microorganisms (Fussmann, Loreau, & Abrams, 2007; Lenski, 2017). The investigation of ecological and evolutionary processes happening in microbial communities facing a disturbance corroborates to better understand features of community stability and bacterial speciation. Moreover, the study of bacterial responses to agrochemical contamination specifically contributes to evidence of anthropogenic effects on the freshwater ecosystems, providing scientific guidance to policymakers.

Ecological responses are defined as changes in population and community dynamics, either directly due to disturbances, resulting from the interaction between species and the environment, or indirectly due to species-species interactions in a biological community (Shade et al., 2012). A variety of features intrinsic to species, such as tolerance to contaminants, or derived from species interactions, such as competition or predation, drive ecological responses following a perturbation event (Clements & Rohr, 2009) and confer stability to a community (Allison & Martiny, 2008). Community stability is comprised of resistance, or the degree to which a community is insensitive to a

disturbance, and by resilience, or the ability to recover after experiencing a disturbance (Shade et al., 2012).

Evolutionary responses are defined as shifts of heritable traits, resulting from new genetic variation or changes in allele frequencies within a population or species. Genetic variation within species arises by recombination or mutation and is fixed through genetic drift by chance in small populations, or by natural selection in larger populations (Nei, Suzuki, & Nozawa, 2010). When a mutation under selection reaches 100% frequency, this is called a selective sweep, which may be restricted to genes (i.e. gene-specific selective sweeps), generally when recombination rates are high relative to the strength of selection, or may happen to complete genomes (i.e. genome-wide selective sweeps) if selection is stronger than recombination (Shapiro & Polz, 2014).

Tracking of ecological and evolutionary changes in microorganisms has been largely improved by culture-independent approaches, particularly next-generation sequencing technologies, such as amplicon and metagenomic sequencing (Shendure & Ji, 2008). Amplicon sequencing of a taxonomic marker, such as the 16S ribosomal RNA (rRNA) gene in prokaryotes, allows a cost-effective description of complex natural microbial communities in environmental samples regarding their taxonomic composition and diversity (Hugerth & Andersson, 2017). Metagenomic sequencing (i.e. the untargeted sequencing of virtually all genomes in a given sample) enables the identification of genes that may be involved directly or indirectly in the response to a perturbation along with the assembly of genomes and tracking allele frequency variation, allowing the assessment of evolutionary responses (Quince, Walker, Simpson, Loman, & Segata, 2017).

Mesocosm studies: simplifying complexity to testing ecological hypothesis

Each ecosystem displays an intrinsic complexity resultant of the many interacting factors: from abiotic, such as temperature and resource availability, to biotic factors, such as species traits and individual variability. The challenge of observational studies in ecology is to disentangle which of the many environmental factors drive species responses while experimental ecology proposes to understand cause-effect relationships by controlling some variables and reducing complexity (Benton, Solan, Travis, & Sait, 2007).

Experiments in ecology vary widely in scale, from small-volume laboratory microcosms (< 1L), large-volume field mesocosms (1 to 10,000 L) to whole-ecosystem manipulations, such as the whole-lakes experiments in the Experimental Lakes Area (Malley & Mills, 1992). Ecotoxicological assessments of agrochemicals are usually conducted in small-scale laboratory assays focusing on the effect of one chemical on a single species. While this approach facilitates understanding the mechanisms of action and identifying species inhibitory concentrations, it ignores potential effects arising from biotic factors, such as the interaction among species, and abiotic factors, such as co-occurring pollutants or climatic conditions (Gessner & Tlili, 2016), which are present in a natural scenario. Field experiments, such as mesocosms or whole-ecosystem manipulations, account for these factors and are therefore complementary to laboratory assays for ecotoxicological assessments (Gessner & Tlili, 2016; M.-P. Hébert et al., 2021).

Outdoor mesocosms combine features of field and laboratory methodologies to estimate causal effects of controlled variables (e.g. concentration of a contaminant) under non-controlled near-natural environmental conditions, such as light exposure, temperature variation and precipitation (Alexander, Luiker, Finley, & Culp, 2016). They are orders of magnitude smaller than a natural ecosystem and usually constrained in duration, limiting their explanatory power, especially when compared to whole-ecosystem experiments (S. Carpenter, 1996; Schindler, 1998). Although mesocosm and microcosm experiments fail to completely replicate a natural ecosystem, they have been historically used to understand ecological processes thanks to their feasibility compared to whole-ecosystem manipulations, which are too expensive, and cause impacts by simulating perturbations in the natural system of interest (Alexander et al., 2016). Mesocosms experiments can be used to explore ecological mechanisms and the drivers of biodiversity loss (Benton et al., 2007) and should be seen as complementary to larger-scale studies.

Thesis structure, objectives and hypotheses

The main goal of this thesis is to provide knowledge on how bacterial communities and populations respond to contamination with agrochemicals in a freshwater mesocosm setup exploring their ecological and evolutionary responses to simulated perturbations. I

used high-throughput sequencing technologies and different bioinformatic approaches to investigate shifts in the composition of species and genes within communities in addition to allelic variation within species through time. Below I summarize the objectives and hypotheses of each of the three studies comprising this thesis.

In the first study (Chapter II), I quantified bacterioplankton community responses to pulse perturbations of two globally relevant pesticides, a glyphosate-based herbicide (GBH) and the neonicotinoid insecticide imidacloprid, in combination with the effect of fertilizers (nutrient pollution). I aimed to investigate ecological responses driven by each pesticide isolated or in combination, under two scenarios of fertilizers input. I described changes in community diversity, species composition, bacterial density and functional diversity through time and across different treatments. My hypothesis was that GBH perturbation would promote direct changes in bacterial composition and reduced richness, as many bacteria have the sensitive class of the glyphosate target enzyme and other naturally occurring species have a resistant enzyme, including the *Agrobacterium* strain whose resistance gene is used in genetically modified crops resistant to glyphosate (Funke, Han, Healy-Fried, Fischer, & Schönbrunn, 2006). On the other hand, I did not expect direct effects on bacteria in the insecticide treatments and I hypothesized that bacterial density would increase as a consequence of the relaxed grazing pressure (M.-P. Hébert et al., 2021). I also predicted that fertilizers treatment would increase bacterial density and that treatments stimulating increased biomass would mask the negative effect of other contaminants (Alexander, Luis, Culp, Baird, & Cessna, 2013). Finally, I expected functional diversity to be more resistant than taxonomic diversity to agrochemical perturbations, as bacterial communities are often redundant for the functions I tested, related to carbon substrate use (Louca et al., 2018).

In the second study (Chapter III), motivated by the fact that some species known to be sensitive to glyphosate were among those favoured by the GBH treatment, I focused on changes that the herbicide promoted in the composition of genes predicted from metagenomic bacterial DNA. More specifically, I tested the hypothesis that antimicrobial resistance genes (ARGs) would explain the success of species facing contamination with GBH. Previous studies with laboratory cultures have shown an association between antimicrobial resistance induced by the presence of glyphosate (Kurenbach et al., 2017;

Kurenbach, Hill, Godsoe, Van Hamelsveld, & Heinemann, 2018) and suggested the role of efflux pumps in the adaptive response to the GBH Roundup (Kurenbach et al., 2015). To our knowledge, only one field study confirmed the correlation between the prevalence of ARGs and the application of glyphosate in soil bacterial communities (Liao et al., 2021) and our study was the first to test it in aquatic microbiomes. We expected to find evidence that, in a complex bacterial community, efflux pumps would explain the adaptive response to glyphosate, as observed in isolated strains in laboratory conditions.

In the third study (Chapter IV), I investigated the evolutionary responses of bacterial populations with similar ecological responses to the GBH treatment. I aimed to test if the ecological responses of successful species observed in Chapter 2 could be explained by the same evolutionary mechanism. I hypothesized that intraspecific diversity in populations surviving to the GBH stress would decrease as a consequence of the selection of resistant individuals. More specifically, I expected to observe genome-wide sweeps, or the fixation of alleles throughout the genome in previously diverse populations, as predicted by the stable-ecotype model, an evolutionary model that explains the origin of distinct phylogenetic groups through occasional diversity purges driven by selection (Cohan, 2001; Cohan & Perry, 2007). Through this chapter, I intended to investigate the generality of the stable-ecotype model. The model has received some support from observational data (Bendall et al., 2016; Jain, Rodriguez-R, Phillippy, Konstantinidis, & Aluru, 2018), but not from experimental data with a controlled selective pressure.

CHAPTER II : Resistance, resilience, and functional redundancy of freshwater bacterioplankton communities facing a gradient of agricultural stressors in a mesocosm experiment

Naíla Barbosa da Costa[1,2], Vincent Fugère[2,3,4,5], Marie-Pier Hébert[2,6], Charles C.Y. Xu[3,6,7], Rowan D.H. Barrett[3,6,7], Beatrix E. Beisner[2,4], Graham Bell[3,6], Viviane Yargeau[8], Gregor F. Fussmann[2,3,6], Andrew Gonzalez[3,6], B. Jesse Shapiro[1,2,3,9,10]

[1] Département des sciences biologiques, Université de Montréal, Montreal, Canada; [2] Groupe de Recherche Interuniversitaire en Limnologie et environnement aquatique (GRIL); [3] Québec Centre for Biodiversity Science (QCBS); [4] Département des sciences biologiques, Université du Québec à Montréal, Montreal, Canada; [5] Département des sciences de l'environnement, Université du Québec à Trois-Rivières, Trois-Rivières, Canada; [6] Department of Biology, McGill University, Montreal, Canada; [7] Redpath Museum, McGill University, Montreal, Canada; [8] Department of Chemical Engineering, McGill University, Montreal, Canada; [9] Department of Microbiology and Immunology, McGill University, Montreal, Canada; [10] McGill Genome Centre, McGill University, Montreal, Canada

Published in *Molecular Ecology*, 30(19): 4771 4788. 10.1111/mec.16100

Minor edits to the published text have been made here, following suggestions by this thesis' reviewers

Abstract

Agricultural pollution with fertilizers and pesticides is a common disturbance to freshwater biodiversity. Bacterioplankton communities are at the base of aquatic food webs, but their responses to these potentially interacting stressors are rarely explored.

To test the extent of resistance and resilience in bacterioplankton communities faced with agricultural stressors, we exposed freshwater mesocosms to single and combined gradients of two commonly used pesticides: the herbicide glyphosate (0-15 mg/L) and the neonicotinoid insecticide imidacloprid (0-60 μg/L), in high or low nutrient backgrounds. Over the 43-day experiment, we tracked variation in bacterial density with flow cytometry, carbon substrate use with Biolog EcoPlates, and taxonomic diversity and composition with environmental 16S rRNA gene amplicon sequencing. We show that only glyphosate (at the highest dose, 15 mg/L), but not imidacloprid, nutrients, or their interactions measurably changed community structure, favoring members of the Proteobacteria including the genus *Agrobacterium*. However, no change in carbon substrate use was detected throughout, suggesting functional redundancy despite taxonomic changes. We further show that communities are resilient at broad, but not fine taxonomic levels: 24 days after glyphosate application the precise amplicon sequence variants do not return, and tend to be replaced by phylogenetically close taxa. We conclude that high doses of glyphosate – but still within commonly acceptable regulatory guidelines – alter freshwater bacterioplankton by favoring a subset of higher taxonomic units (i.e. genus to phylum) that transiently thrive in the presence of glyphosate. Longer-term impacts of glyphosate at finer taxonomic resolution merit further investigation.

Introduction

Agricultural expansion and intensification are major drivers of global environmental change in both terrestrial and aquatic ecosystems (Song et al., 2018; Springmann et al., 2018; Tilman et al., 2001). Chemicals derived from agricultural landscapes, such as fertilizers and pesticides, are among the main sources of freshwater pollution (Vörösmarty et al., 2010), leading to eutrophication (S. R. Carpenter et al., 1998; Keatley, Bennett, Macdonald, Taranu, & Gregory-Eaves, 2011) and biodiversity loss (DeLorenzo, Scott, & Ross, 2001; Relyea, 2009; Stehle & Schulz, 2015). Anthropogenic climate change may

intensify these effects as variation in precipitation patterns and increased temperatures affect agrochemicals fate, transport, and behavior in surface and groundwater (Bloomfield, Williams, Gooddy, Cape, & Guha, 2006; Jeppesen et al., 2009). Agricultural runoff to waterbodies particularly increases after storms, acting as a pulse perturbation (Cedergreen & Rasmussen, 2017) while bringing a mixture of nutrients, herbicides and insecticides that may interact to affect aquatic microbial taxa (Flood & Burkholder, 2018) and communities (Lozano & Pratt, 1994; Starr, Bargu, Maiti, & DeLaune, 2017). The impact of agricultural contaminants may depend on whether they are applied alone or in combination (Altenburger et al., 2013), and the effects of combinations may be difficult to predict based upon data from single contaminants, possibly due to complex interactions within diverse bacterial communities (Romero, Acuña, & Sabater, 2020).

Agricultural activity has a major impact on bacterioplankton (Kraemer et al., 2020) and, as a consequence, on the ecosystem processes they provide; e.g. decomposition of organic matter (Piggott, Niyogi, Townsend, & Matthaei, 2015) and nutrient cycling (Romero et al., 2020). Altering these processes may have broad consequences for aquatic ecosystem productivity, food webs, and the human activities that depend upon them (S. R. Carpenter, Stanley, & Vander Zanden, 2011).

Nutrient pollution is among the most important stressors affecting biodiversity in lakes (Birk et al., 2020). It promotes eutrophication (V. H. Smith, Joye, & Howarth, 2006), which can increase bacterial biomass, reduce phytoplankton diversity, and trigger harmful algal blooms (Paerl, Otten, & Kudela, 2018; V. H. Smith & Schindler, 2009). While few studies have addressed individual and combined effects of fertilizers with herbicides or insecticides on phytoplankton and zooplankton communities (Baker, Mudge, Thompson, Houlahan, & Kidd, 2016; Chará-Serna, Epele, Morrissey, & Richardson, 2019; Geyer et al., 2016), analogous assessments of bacterioplankton are more scarce. Yet, similar to other planktonic communities, bacteria may also be directly or indirectly (e.g. through trophic effects) affected by the individual or combined effects of these stressors, despite not being their intended targets (Muturi, Donthu, Fields, Moise, & Kim, 2017).

The herbicide glyphosate, mainly formulated commercially as Roundup, and the neonicotinoid insecticide imidacloprid (available in different commercial formulations) are among the most commonly used pesticides worldwide (Benbrook, 2016; Simon-Delso et

al., 2015), despite restrictions on their use in different jurisdictions. In North America and the European Union, common benchmarks to protect aquatic life range from 800 to 26,600 μg/L of glyphosate for long-term (chronic) exposure, and between 27,000 to 49,000 μg/L for short-term (acute) exposure (CCME, 2012; EFSA, 2016; EPA, 2019). In contrast, lower concentrations of imidacloprid are considered safe for aquatic invertebrates, ranging from 0.009-0.385 μg/L (CCME, 2007; EFSA, 2014; EPA, 2019) (Table II.S1). Most of these criteria were developed based on toxicity tests on individual eukaryotic organisms, and it remains unclear how bacterial communities respond to these concentrations considered "safe for aquatic life" and what consequences their responses might have on the ecosystem functions they provide.

Glyphosate is a broad-spectrum synthetic phosphonate herbicide used for weed control. It acts by inhibiting the enzyme enolpyruvylshikimate-3-phosphate synthase (EPSPS) involved in the biosynthesis of aromatic amino acids essential to plants, its target group, but also to many fungi and bacteria (Pollegioni, Schonbrunn, & Siehl, 2011). However, some microorganisms are resistant to glyphosate either by expressing an insensitive form of the target enzyme (Funke et al., 2006; Healy-Fried, Funke, Priestman, Han, & Scho, 2007) or by metabolizing the molecule and using it as a phosphorus source (Hove-Jensen, Zechel, & Jochimsen, 2014). Glyphosate could therefore select for resistant species within bacterial communities (Muturi et al., 2017). Moreover, as it may prevent the growth of some phytoplankton species (Smedbol, Lucotte, Labrecque, Lepage, & Juneau, 2017), bacterioplankton could be affected indirectly, for example by reduced competition with phytoplankton.

Unlike glyphosate, imidacloprid is an insecticide commonly used as a seed-coating agent intended to control sapling damage from piercing-sucking insects (CCME, 2007; Jeschke & Nauen, 2008). It acts on insect nervous systems (Roberts & Hutson, 1999) and can be toxic to many aquatic invertebrates, especially insects and crustaceans (Morrissey et al., 2015). Although it is not known to inhibit bacteria directly, it could affect them indirectly via trophic effects on their predators or grazers. If imidacloprid reduces total zooplankton biomass, for example, a reduction in predation pressure could promote an increase in bacterioplankton biomass. Ecosystem functions provided by

bacterioplankton, such as carbon use, could subsequently be affected, as has been observed in experiments with other insecticides (Thompson et al., 2016).

Bacterioplankton are important drivers of energy and nutrient cycling in freshwater ecosystems (Falkowski et al., 2008; Konopka, 2009), and more observations are needed to understand how they respond to anthropogenic disturbances (Allison & Martiny, 2008). They may respond with detectable changes in species composition (Allison & Martiny, 2008) that could be permanent, thereby providing a measure of the historical impact of anthropogenic activities on ecosystem health (Kraemer et al., 2020). Alternatively, community composition could be resistant or resilient to changes (Shade et al., 2012). Even if disturbances alter community composition, ecosystem processes may remain stable if pre- and post-disturbance communities are functionally redundant (Allison & Martiny, 2008).

Functional redundancy is thought to be common in microbial communities, as most metabolic pathways controlling biogeochemical cycles are encoded by several different phylogenetic groups. Certain functions, such as photosynthesis and methanogenesis, are however phylogenetically restricted (Falkowski et al., 2008). It is likely that communities are partially redundant for general functions like respiration or biomass production, but non-redundant for more specific functions encoded by unique taxa (Louca et al., 2018). The prevalence of, and reasons for microbial community resistance, resilience, and functional redundancy are still debated (Allison & Martiny, 2008; Shade et al., 2012), particularly in response to novel anthropogenic disturbances which increasingly involve combinations of stressors.

In this study, we experimentally tested the effects of pulse applications of glyphosate and imidacloprid, under low (mesotrophic) or high (eutrophic) nutrient conditions, on bacterioplankton community density, taxonomic composition and richness, and functions related to carbon substrate use. To do so, we filled 1,000 L mesocosms with water and planktonic organisms from a pristine lake located on a mountaintop of a protected area with no history of agricultural activity. Using a regression design, we applied gradients of pesticide concentrations (Figure II.1), spanning ranges observed in surface runoff and freshwater systems (Hénault-Ethier et al., 2017; Morrissey et al., 2015; van Bruggen et al., 2018). Highest doses applied are considered harmful to eukaryotic

organisms upon which nationwide water quality guidelines are based (Table II.S1). To quantify individual and interactive effects of agricultural stressors, we applied these pesticides alone and in combination, and in the presence or absence of nutrient enrichment simulating fertilizer pollution. Pesticides were applied as pulse perturbations to mimic how these contaminants reach natural freshwater ecosystems from agricultural fields, while nutrient enrichments were applied as press treatments to mimic mesotrophic and eutrophic conditions.

We hypothesize that glyphosate will change bacterial community composition and reduce richness, as many taxa depend on the target enzyme (EPSPS) to synthesize aromatic amino acids, while other species encode a resistant allele of EPSPS (Funke et al., 2006; Healy-Fried et al., 2007; Rainio et al., 2021) or are able to metabolize glyphosate (Hove-Jensen et al., 2014). While imidacloprid is less likely to directly impact bacteria, we hypothesize that it can exert indirect effects due to its potential toxicity to aquatic invertebrates (Chará-Serna et al., 2019), releasing grazing pressure on bacterial communities and increasing their density. When applied in combination with glyphosate, imidacloprid may therefore delay or mask the effects of glyphosate on bacterioplankton community structure. Similarly, fertilizers might also increase microbial productivity and mask negative effects of glyphosate, as it does with other contaminants (Alexander et al., 2013). We also expect some positive effects of glyphosate on bacterial density, as it may serve as a source of phosphorus for some species (Hove-Jensen et al., 2014). Finally, we predict that functional diversity will be less prone to changes than taxonomic composition, as bacterial communities tend to be functionally redundant (Louca et al., 2018). We thus expect to detect changes in bacterial community composition and species richness at lower pesticide doses, and changes in functional diversity only at higher doses, or not at all.

Materials and methods

Experimental design and sampling

We conducted a mesocosm experiment at the Large Experimental Array of Ponds (LEAP) platform at McGill University's Gault Nature Reserve (45°32'N, 73°08'W), a protected area with no history of agricultural pollution (Beauséjour, Handa, Lechowicz,

Gilbert, & Vellend, 2015) in Quebec, Canada. The pond mesocosms at LEAP are connected to a reservoir that receives water from the upstream Lake Hertel through a 1 km pipe by gravity. On May 11th, 2016 (99 days prior to the start of the experiment), 100 ponds were simultaneously filled with 1,000 L of lake water, to acclimate communities to the mesocosm setting. When filling ponds we used a coarse sieve to prevent fish introduction. To maximize initial homogeneity among communities (before treatments), and because this study focuses on planktonic microbes, no sediment substrate was added to the ponds. Tadpoles and large debris such as leaves and pollen were periodically removed with a net before the experiment commenced. Additional lake water was added on a biweekly basis (~10% of total volume) between May and August to ensure a continuous input of lake bacterioplankton, tracking seasonal changes in the source lake community, and to homogenize communities across ponds. The experiment reported here used 48 of these pond mesocosms from August 17th (day 1) to September 28th (day 43), and it is part of a collaborative experiment that also assessed responses of zooplankton in the same set of ponds (M.-P. Hébert et al., 2021) and phytoplankton responses in a subset of these ponds for a longer period of time (Fugère et al., 2020).

Throughout our experiment, phosphorus (P) and nitrogen (N) were simultaneously added biweekly to simulate nutrient enrichment at a constant rate, starting on August 10th, 7 days before the first sampling day to ensure communities would have passed their exponential growth phase before the first pulse of pesticides was applied. Our nutrient treatment included two levels, with target concentrations of 15 µg P/L (hereafter referred as low-nutrient treatment) typical of mesotrophic Lake Hertel (Thibodeau, Walsh, & Beisner, 2015) and 60 µg P/L (high-nutrient treatment; eutrophic conditions). Nutrient solutions were made using nitrate (KNO_3) and phosphate (KH_2PO_4 and K_2PO_4) preserving the same N:P molar ratio (33:1) found in Lake Hertel; the target concentrations were therefore 231 µg N/L and 924 µg N/L for low- and high- nutrient treatments respectively. Over the course of the experiment, the average total P (TP) concentration measured in the source lake was 20.4 µg/L (standard error, SE=±1.3) and the average TP achieved in ponds with no pesticide addition was 13.6 µg/L (SE=±0.71) and 36.7 µg/L (SE=±10.8) respectively for low- and high-nutrient treatment. The average total N (TN)

concentration was 556.9 µg/L (SE=±60.7) at Lake Hertel, 407.8 µg/L (SE=±32.7) and 789.0 µg/L (SE=±177.6) respectively in control ponds with low and high nutrient inputs.

Within each nutrient treatment, ponds received varying amounts of the herbicide glyphosate or the insecticide imidacloprid, separately or in combination, in a regression design with seven levels of pesticide concentration plus controls with no pesticide addition (Figure II.1A). The seven levels of target concentration were: 0.04, 0.1, 0.3, 0.7, 2, 5.5 and 15 mg/L for glyphosate and 0.15, 0.4, 1, 3, 8, 22 and 60 µg/L for imidacloprid. There was no replication for each combination of nutrient and pesticide concentration, which is compensated by the wide gradient of pesticides concentration established in the regression design (Figure II.1A). Glyphosate was added in the form of Roundup Super Concentrate (Monsanto©) and target concentrations calculated based on its glyphosate acid content, while imidacloprid was added in the form of a solution prepared with pure imidacloprid powder (Sigma-Aldrich, Oakville, Canada) dissolved in ultrapure water. Treatment ponds received two pulses of pesticides (at days 6 and 34 of the experiment) while nutrients were applied biweekly to maintain a press treatment. The target concentrations of glyphosate and imidacloprid were well correlated with the measured concentrations in the ponds (Figure II.S1A-B) with the exception of a few ponds receiving the highest imidacloprid dose which reached lower concentrations than intended, especially after the second pulse (Figure II.S1C). These ponds nonetheless reached higher concentrations than ponds lower on the imidacloprid gradients (i.e., a clear gradient was established).

Bacterioplankton communities were sampled at six different timepoints (Figure II.1B): one before pesticide application (day 1); three between pulse 1 and pulse 2 applications (days 7, 15, and 30); and two timepoints after the second pulse (days 35 and 43). Pesticide quantification was performed immediately after each pulse application (days 6 and 34) and at two time points between them (days 14 and 29) while nutrients were quantified on the same days as bacterioplankton except for days one and 35 (Figure II.1B).

Water samples for nutrient and microbial community analyses were collected from each mesocosm with integrated samplers (made of 2.5 cm-wide PVC tubing) and stored in dark clean 1L Nalgene (Thermo Scientific) bottles triple-washed with pond water. To

avoid cross contamination, we sampled each pond with a separate sampler and bottle. We kept bottles in coolers while sampling and then moved them to an on-site laboratory, where they were stored at 4 °C until processing, for no longer than 4 hours. Water samples for pesticide quantification were collected immediately after pesticide application (days 6 and 34) in a subset of ponds englobing each gradient and in a smaller subset between the pulses (days 14 and 29). They were stored in clear Nalgene bottles (1 L), acidified to a pH < 3 with sulfuric acid and frozen at -20 ºC until analysis.

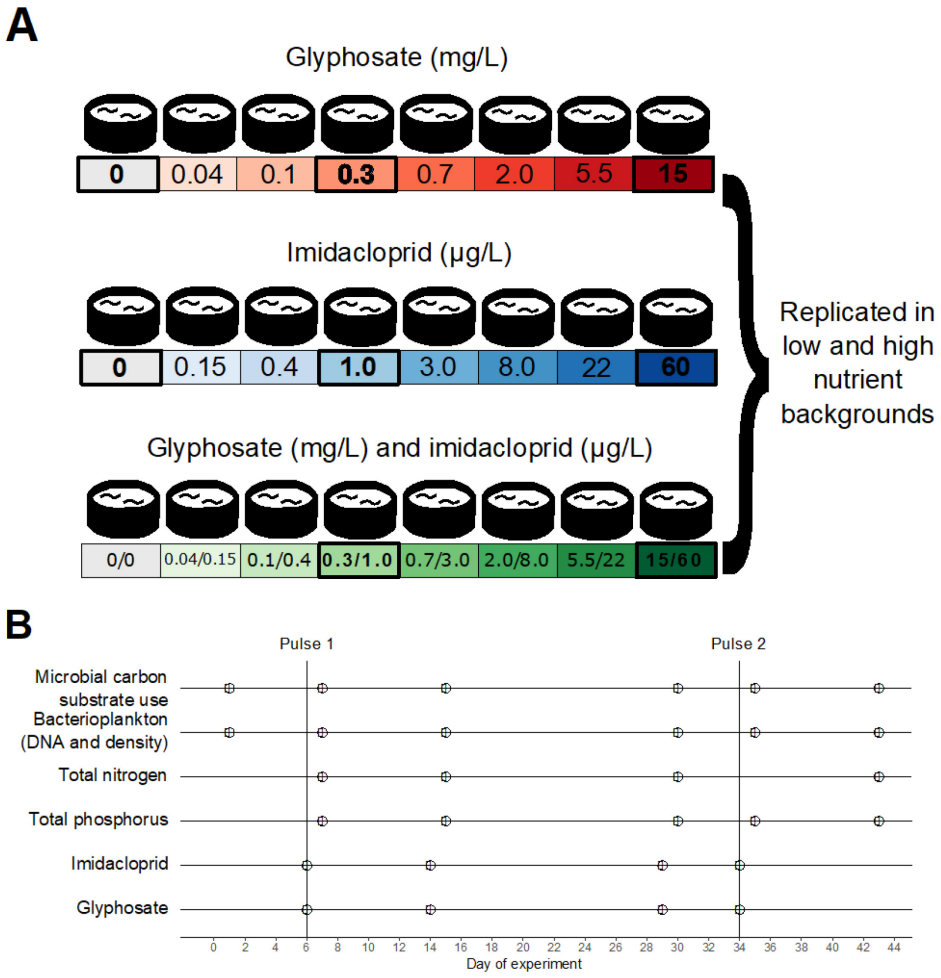


Figure II.1 Experimental design and sampling timeline A) In total, 48 mesocosms (ponds) at the Large Experimental Array of Ponds (LEAP) at the Gault Nature Reserve were filled with 1,000 L of pristine lake water and received two pulses of the pesticides glyphosate and imidacloprid, alone or in combination, at two different nutrient enrichment scenarios. Each box represents an experimental pond and those outlined in bold indicate ponds sampled for DNA extraction and 16S rRNA gene amplicon sequencing. B) The experiment lasted 43 days and pesticides were applied on days 6 (pulse 1) and 34 (pulse 2). Dates of sampling for each variable are indicated with points. Nutrients were added every two weeks at a constant dose, starting seven days before the first sampling day.

Nutrient and pesticide quantification

Quantification of TP and TN from unfiltered water samples were processed at the GRIL (Interuniversity Group in Limnology) analytical laboratory at the Université du Québec à Montréal following standard protocols as outlined by McComb (2002). Duplicate subsamples (40 mL) of water sampled from each pond were stored in acid-washed glass tubes and kept at 4 ºC until nutrient concentrations were quantified. TN concentration was determined using the alkaline persulfate digestion method coupled with a cadmium reactor (Patton & Kryskalla, 2003) in a continuous flow analyzer (OI Analytical, College Station, TX, USA). TP was estimated based on optical density in a spectrophotometer (Biocrom Ultrospec 2100pro, Holliston, MA, USA) after persulfate digestion through the molybdenum blue method (Wetzel & Likens, 2000). Glyphosate and imidacloprid concentrations were quantified through liquid chromatography coupled to mass spectrometry using an Accela 600-Orbitrap LTQ XL (LC–HRMS, Thermo Scientific). The method consisted of heated electrospray ionization (HESI) in negative mode for glyphosate, acquisition in full scan mode (50-300 m/z) at high resolution (FTMS = 30,000 m/Dz) and the same LC-HRMS system but using positive HESI mode for imidacloprid (mass range 50-700m/z). Limits of detection were 1.23 and 1.44 µg/L for glyphosate and imidacloprid respectively, while quantification thresholds were respectively 4.06 µg/L, and 4.81 µg/L. Samples falling below limits of detection were pre-concentrated with a factor of 40X (10 mL samples were reconstituted to 250 µL) and their final concentration were back-calculated according to the concentration factor.

Estimating bacterial density through flow cytometry

To estimate the density of bacterial cells, we fixed 1 mL of the 1 L sampled pond water with glutaraldehyde (1% final concentration) and flash froze this subsample in liquid nitrogen (Gasol & Del Giorgio, 2000; Ruiz-González et al., 2018). We stored samples at -80 ºC until they were processed via a BD Accuri C6 flow cytometer (BD Biosciences, San Jose, CA, USA). After samples were thawed at room temperature (18-20 °C), we prepared dilutions (1:25) with Tris-EDTA buffer (Tris-HCl 10 mM; EDTA 1 mM; pH 8) and aliquoted in two duplicate tubes. Samples were then stained with Syto13 Green-Fluorescent Nucleic Acid Stain (0.1 v/v in DMSO; ThermoFisher S7575) and incubated in

the dark at room temperature (18-20 °C) for 10 min. To validate the equipment calibration, we ran BD TruCount Absolute Count Tubes (BD Biosciences) each day, prior to sample processing. Samples were run until reaching 20,000 events, at a rate of 100-1,000 events/s in slow fluidics (14 µL/min). Events within a predefined gate on a 90° light side scatter (SSC-H) versus green fluorescence (FL1-H) cytogram were used for cell counts estimation. This inclusive gate was defined to maximize cell counts accuracy by excluding background noise and large debris. Bacterial density was estimated based on cell counts detected within the gate, flow volume, and sample dilution. We calculated the average bacterial density for each pair of analytical duplicates with a coefficient of variation (CV, i.e., ratio between the standard deviation and average of the duplicate values) less than 0.08. If the CV was greater than or equal to 0.08, the sample was run a third time, and the outlying value was discarded before taking the mean of the two remaining samples.

Carbon substrate utilization patterns

We used Biolog EcoPlate® assays (Hayward, CA, USA) to infer community-level utilization of dissolved organic carbon by microbes. For all treatments (Figure II.1A) and at each of the six sampled timepoints (Figure II.1B), we added 125 µL of unfiltered pond water to each well of the EcoPlates. Each plate contains, in triplicates, 31 different organic carbon substrates and water controls. These substrates can be grouped into five main guilds (amines/amides, amino acids, carbohydrates, carboxylic acetic acids and polymers), as summarized in Table II.S2. We measured the optical density at 590 nm in each well as a proxy for microbial carbon substrate use, since it causes a concomitant reduction of the redox-sensitive tetrazolium dye, whose color intensity is measurable at this wavelength. Plates were incubated in the dark at room temperature (18-20 °C) and well absorbance was measured daily until an asymptote was reached (Ruiz-González et al., 2018; Ruiz-González, Niño-García, Lapierre, & del Giorgio, 2015). For each daily measurement, an average well color development (AWCD) was calculated. To correct for variation in inoculum density we selected substrate absorbance values of the plate measurements with AWCD closest to 0.5 (usually after 3-8 days of incubation) as suggested in Garland (2001). We then calculated the blank-corrected median absorbance of each substrate at each sampled timepoint for analyses.

DNA extraction, 16S rRNA gene amplification and sequencing

We selected a subset of ponds for DNA extraction and subsequent analyses (outlined in bold in Figure II.1A) to assess bacterioplankton community responses at the extremes and the middle of the experimental gradient. From each timepoint and nutrient treatment, we chose two control ponds (beginning of the gradient, no pesticide addition), ponds with the third lowest concentration (middle of the gradient) of each or both pesticides (1 µg/L imidacloprid and/or 0.3 mg/L glyphosate), and ponds with the highest concentration (end of the gradient) used in the experiment for each or both pesticides (60 µg/L imidacloprid and/or 15 mg/L glyphosate). We selected ponds with high concentrations of pesticides to maximize the chance of detecting a response from the bacterial community. That said, we still kept concentrations that fall below available regulatory acceptable concentrations for glyphosate in North America (Table II.S1), allowing us to ask whether changes in bacterial communities can be detected at concentrations considered safe for aquatic eukaryotes in a region where glyphosate is extensively used (Benbrook, 2016; Simon-Delso et al., 2015). In total, we sampled 16 of the 48 experimental ponds at six timepoints, yielding a total of 96 samples for 16S rRNA amplicon sequencing (Figure II.1B). After sampling 1 L of pond water as described above, we immediately filtered 250 mL through a 0.22 µm pore size Millipore hydrophilic polyethersulfone membrane of 47 mm diameter (Sigma-Aldrich, St. Louis, USA) and stored filters at -80 ºC until DNA extraction. We extracted and purified total genomic DNA from frozen filters using the PowerWater DNA Isolation Kit (MoBio Technologies Inc., Vancouver, Canada) following the manufacturer's protocol, that includes a 5-min vortex agitation of the filter with beads and lysis buffer to enhance cell lysis. We quantified genomic DNA with a Qubit 2.0 fluorometer (ThermoFisher, Waltham, MA, USA) and used 10 ng to prepare amplicon libraries for paired-end sequencing (2 x 250 bp) on two Illumina MiSeq (Illumina, San Diego, CA, USA) runs. We performed a two-step polymerase chain reaction (PCR) targeting the V4 region of the 16S rRNA gene, with primers U515_F and E786_R, as described in Preheim et al. (2013). Further details on PCR reactions, library preparation and amplicon sequencing, including positive controls (mock communities) and negative controls are described in the Supplementary Material.

Sequence data processing

We used idemp (https://github.com/yhwu/idemp) to demultiplex barcoded fastq files from the sequencing data, and cutadapt to remove remaining Illumina adapters (Martin, 2011). The DADA2 package (Callahan et al., 2016) in R was used to filter and trim reads, using the default filtering parameters with a maximum expected error (maxEE) score of two. Reads were trimmed on the left to remove primers and those shorter than 200 or 150 bp were discarded, respectively, for forward and reverse reads. DADA2 was also used to infer amplicon sequence variants (ASVs), remove chimeras and finally obtain a matrix of ASV counts in each sample for each MiSeq run independently. We used the default parameters of the "learning error rates" function with the multithread option enabled. The number of raw reads and non-chimeric reads obtained from each sample are summarized in Table II.S3 (average raw reads per sample: 43,159; SE=2,245). Excluding mock communities, extraction blanks and PCR controls, we obtained 1,787,412 raw reads in the first run and 4,702,355 in the second run, of which we retained, respectively, 1,565,021 and 4,188,644 non-chimeric reads. PCR negative controls and extraction blanks produced 214 non-chimeric reads in total; these were excluded from downstream analyses as we only included samples with a minimum of 6,000 reads. Of the 30 expected sequences from the custom mock community (Preheim et al., 2013), DADA2 found 25 exact sequence matches, producing 5 false negatives and 7 false positives (for a total of 32 sequences). In the ATCC mock, 23 of the 24 expected sequences were found, with only 1 false negative but 10 false positives (for a total of 33 sequences). False positives were closely related to the expected sequence match and, compared to sequence clustering-based methods, ASV inference showed a better performance. We concatenated DADA2 abundance matrices from each MiSeq run and then used TaxAss (Rohwer, Hamilton, Newton, & McMahon, 2018) to assign ASV taxonomy with a database specifically curated for freshwater bacterioplankton, FreshTrain (Newton, Jones, Eiler, Mcmahon, & Bertilsson, 2011), and GreenGenes (DeSantis et al., 2006), with a minimum bootstrap support of 80% and 50%, respectively. After performing a multiple sequence alignment with the R package DECIPHER (Wright, 2016), we constructed a maximum likelihood phylogenetic tree using the phangorn package following recommendations made by Callahan (2016). For subsequent

analyses, we imported the ASV abundance matrix together with taxonomic assignments and environmental data as an object in the phyloseq package (McMurdie & Holmes, 2013) in R. We removed sequence data identified as mitochondria or chloroplast DNA and normalized read counts using the DESeq2 package (Love, Huber, & Anders, 2014), which performs a variance stabilizing transformation without discarding reads or samples (McMurdie & Holmes, 2014), which is important in the context of high read depth variation, as observed among our samples (Table II.S3). As normalizations such as the DESeq2 method tend to reduce the importance of dominant taxa while inflating the importance of rare taxa (McKnight et al., 2019), for comparison with DESeq2, we additionally normalized the abundance matrix in two ways: (1) by calculating relative abundances (proportions) of each ASV, and (2) by rarefying to 10,000 reads (948 ASVs and 7 samples were consequently removed). These two alternative normalizations are presented in the Supplementary Materials, and are generally concordant with the DESeq2 results in the main text. For most compositional analyses in the main text, we calculated the estimated absolute abundance (EAA) of ASVs per sample by multiplying the DESeq2 normalized ASVs relative abundance by the total bacterial cell counts found in the sample through flow cytometry (Z. Zhang et al., 2017).

Statistical analyses

To assess resistance and resilience to experimental treatments, we compared changes in bacterial community density, microbial carbon substrate use, as well as bacterioplankton community taxonomic structure (richness and composition), as explained in detail below. We conducted all statistical analyses in R version 3.5.1 (R Core Team, 2020). As we tested hypotheses of different treatment effects at different timepoints, we applied a Bonferroni correction for multiple hypothesis testing.

Treatment effects on bacterioplankton density

Time series of bacterial density were analyzed with a generalized additive mixed model (GAMM) with the mgcv R package (Wood, 2017) to quantify the singular and interactive effects of nutrient and of each pesticide treatment on bacterioplankton density as a function of time while accounting for nonlinear relationships. Glyphosate and imidacloprid target concentrations were rescaled (from 0 to 1) to match the scale of the nutrient treatment factor (binary) and we tested for their effect individually or in

combination. Individual mesocosms (ponds) were included as a random effect (random smooth) to account for non-independence among measurements from the same pond over time. Model validation was performed by investigating residual distributions, comparing fitted and observed values and checking if basis dimensions (k) of smooth terms were not too low. The model fit (adjusted R^2) and further details on predictors used in the model, including their statistical significance, are provided in Table II.S4.

Treatment effects on carbon substrate use

We quantified treatment effects on the number of carbon substrates used at each pond and timepoint with a GAMM with the same terms as the GAMM described above for modeling bacterial density. More details are provided in Table II.S4. To assess the effects of the treatments on carbon substrate utilization patterns by microbial communities over time, we built principal response curves (PRCs) (Auber, Travers-Trolet, Villanueva, & Ernande, 2017). PRCs are a special case of partial redundancy analysis (pRDA) in which time and treatments, expressed as ordered factors, are used as explanatory variables, while community composition is the multivariate response. Time is considered as a covariable (or conditioning variable) whose effect is partialled out, and changes in community composition with the treatments over time are always expressed as deviations from the control pond at each timepoint. PRCs also assess the contribution of each species to the treatment effect through the taxa weight (also known as species score) usually displayed in the right y-axis of a PRC diagram (Van den Brink, den Besten, bij de Vaate, & ter Braak, 2009). The significance of the first PRC axis was inferred by permuting the treatment label of each pond while keeping the temporal order, using the permute R package (Simpson, 2019) followed by a permutation test with the vegan R package (A. J. Oksanen et al., 2018). Before performing PRCs we transformed the community matrix (containing carbon substrate use data) using the Hellinger transformation (Legendre & Gallagher, 2001). PRC of community carbon utilization patterns was performed for the 31 substrates individually and grouped into five guilds (Table II.S2).

Treatment effects on bacterioplankton community taxonomic structure

To infer the impact of treatments on bacterioplankton taxonomic diversity over time, we calculated alpha diversity as richness (number of observed ASVs) and as the exponent of the Shannon index (or Hill numbers (Jost, 2006)) of each sample after

rarefying the ASV abundance matrix to 10,000 reads without replacement and modelled their response to pesticide and nutrient treatments using GAMMs. Model equations, their fit (adjusted R^2) and statistics of significant terms are reported in Table II.S5. In this analysis, pesticide treatments were considered factors (low vs. high) because 16S rRNA reads data were only available for a subset of concentrations (Figure II.1A). Pesticide and nutrient treatments were coded as ordered factors and models were validated after investigation of residual distributions, comparison of fitted and observed values and checking if the basis dimension (k) of smooth terms was sufficiently large.

To assess differences in community composition, we calculated weighted UniFrac distances (Lozupone & Knight, 2005) and Jensen-Shannon divergence (JSD) among the subset of samples selected for DNA analyses and represented them in principal coordinate analysis (PCoA) bidimensional plots. These two metrics are complementary as the first is weighted for phylogenetic branch lengths unique to a particular treatment, and the second assesses changes in community composition at the finest possible resolution, tracking ASVs regardless of their phylogenetic relatedness. We performed a series of permutational analyses of variance (PERMANOVA) based on weighted UniFrac distances and JSD to test the effect of treatments (as factors) on community composition at four sampled timepoints separately: at day 1 (before any treatment was applied), day 7 (immediately after the first pulse), day 15 (11 days after the first pulse), day 30 (immediately before the second pulse) and day 43 (last day of the experiment, after the second pulse). We also performed an analysis of multivariate homogeneity (PERMDISP) to test for homoscedasticity in groups dispersions (Anderson, 2006) because the PERMANOVA may be sensitive to non-homogeneous dispersions within groups and thus mistake it as among-group variation (Anderson, 2001). A significant PERMDISP ($p<0.05$) indicates different within-group dispersions and thus should be used in combination with visual inspection of the ordination plots to interpret the PERMANOVA results.

Using EAA after read depth normalization with DESeq2, we further visualized bacterioplankton community temporal shifts with PRCs, asking if the extent of community turnover varied across phylogenetic levels. Separate models were built for ASVs grouped at various phylogenetic levels, from phylum to genus. For each PRC model, we evaluated the proportion of variance (inertia) explained by the conditional variable (time) and the

constrained variable (treatments), as well as the proportion of explained variance per axis (the eigenvalue of each RDA axis divided by the sum of all eigenvalues). We used these values to decide which PRC model, if at the phylum, class, order, family, genus or ASV level, best explained the variation in the data, and we tested for the significance of the first PRC axis through a permutation test with the permute and vegan packages in R (A. J. Oksanen et al., 2018; Simpson, 2019). Taxa weights representing the affinity of the most responsive taxa with the treatment response curve are displayed the right y-axis of each PRC diagram. Before performing each PRC we transformed the community matrix using the Hellinger transformation (Legendre & Gallagher, 2001).

The abundance of the three genera with the highest PRC taxa weights were modeled with GAMMs to explore how treatments impacted their (potentially non-linear) abundances over time, and to provide further validation of the treatment effects detected by PRCs. The GAMM response variable was the log-transformed (log (1+x), where x is the variable) EAA of each of the three genera, after reads had been rarefied to 10,000 reads per sample without replacement. We opted for using rarefied data instead of DESeq2 normalization which is intended for community analyses (Weiss et al., 2017) and the GAMMs focused on specific taxa of interest. Modeled abundances were visualized with the R package itsadug (Van Rij, Wieling, Baayen, & van Rijn, 2020).

Results

Bacterial cell density is weakly affected by glyphosate while microbial community carbon substrate utilization is resistant to all stressors

Overall bacterial cell density showed a strong but non-linear increase over time across all ponds (GAMM, effect of time: F=17.5, p<0.001, Table II.S4; Figure II.2). The time-independent effect of nutrients on bacterial cell density was weak but positive (GAMM, t=4.1, p<0.001), and, over time, glyphosate had a weak positive effect on bacterial density (GAMM, factor-smooth interaction between time and glyphosate: F=6.6, p<0.001) (Table II.S4). The interactive effect of nutrients and glyphosate was also weak, and not significant after Bonferroni correction for multiple testing (GAMM, F=5.7, uncorrected p=0.018, Table II.S4). Overall, these results indicate that, despite increasing

over time across ponds, bacterioplankton densities also slightly increased in response to nutrient and glyphosate addition.

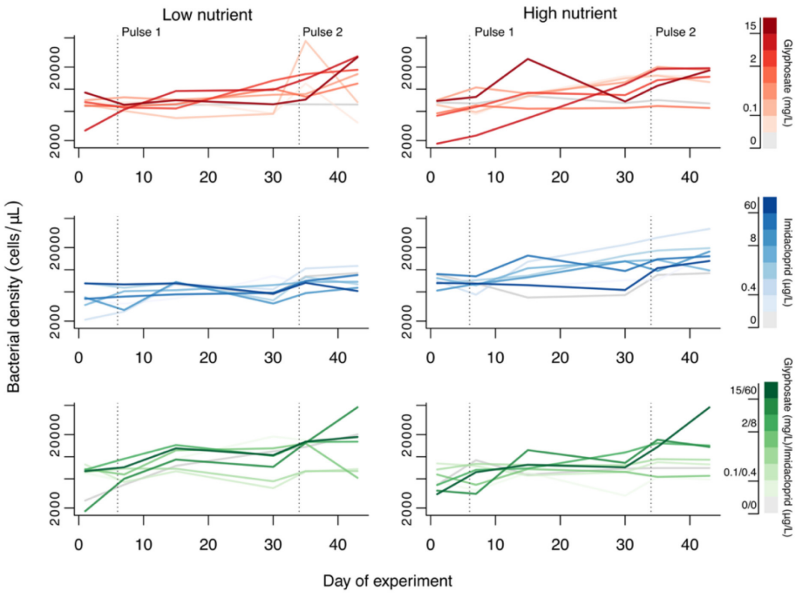


Figure II.2 Bacterial cell density dynamics during the experiment. Total bacterial density is plotted over time in ponds with low- or high- nutrient enrichment. Dashed vertical lines indicate days of pesticide pulses application. Ponds with both glyphosate and imidacloprid follow the same gradient pattern as treatments with either of these pesticides applied alone.

The number of carbon substrates used by the microbial community diminished slightly over time (GAMM, F=6.0, p<0.001, Table II.S4). However, neither glyphosate, imidacloprid, nutrients, nor their interactions had significant effects on carbon substrate utilization as assayed by EcoPlates (Table II.S4). In addition, the PRC analysis did not reveal any significant treatment effects on microbial utilization of any of the 31 unique carbon substrates when considered separately (Figure II.3A; permutation test for the first constrained eigenvalue, F=12.28 p=0.295) or when grouped into guilds (Figure II.3B; F=34.46 p=0.355). To simplify visualization and facilitate comparison with treatments selected for community taxonomic characterization, the PRCs in Figure II.3 included the same ponds as those used for DNA analyses. PRCs including all ponds in the tested gradient showed similar results (Figure II.S2A and Figure II.S2B). We conclude that, despite slight changes in the number of substrates being used over time, none of the treatments significantly affected microbial community-level carbon utilization profiles.

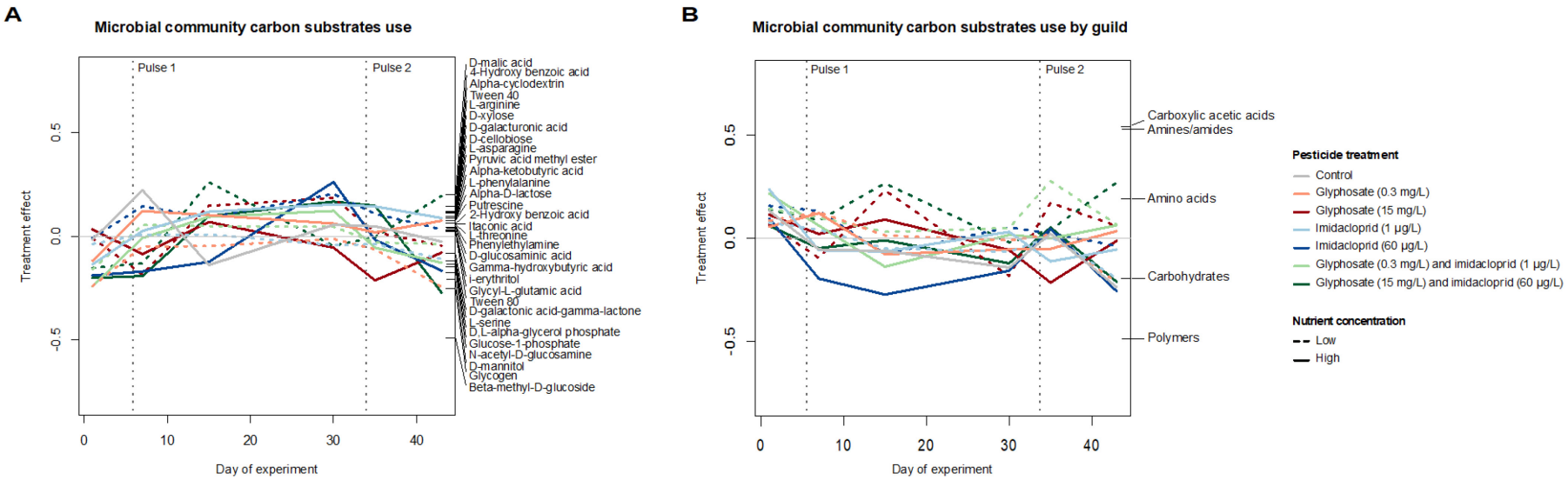


Figure II.3 Microbial community carbon substrate utilization. Principal response curves (PRCs) of selected experimental treatments show no significant difference between controls and pesticide treatments when microbial communities are described according to (A) their ability to metabolize 31 different carbon substrates when analysed individually or (B) when grouped into guilds. Weights of each tested compound or guild are shown along the Y axis (right) and respect the same scale of the right Y axis. Dashed vertical lines indicate days of pesticides pulses application. For ease of comparison, the PRCs were calculated based on the subset of samples for which DNA was extracted. The PRC displayed in (A) explains 15.1% of the variation while the one displayed in (B) explains 42.2%, suggesting that grouping substrates into guilds improves the explanatory power of the PRC.

Bacterioplankton community structure responses

Glyphosate has a minor time-independent effect on community diversity and a major effect on community composition over time

We calculated two metrics of bacterioplankton community alpha diversity in each sample: taxon richness, estimated as the logarithm of the total number of observed ASVs after rarefying (Figure II.4A), and the exponent of the Shannon index, which combines information about ASV richness and evenness (Figure II.4B). No significant time-dependent effect of any treatment was detected, although ponds with high glyphosate concentration (15 mg/L) had a lower Shannon diversity when averaged across all timepoints (GAMM, t=-3.51, p=0.001, Table II.S5), and the same was observed for ASV richness but with a non-significant effect after multiple test correction (GAMM, t=-2.89, uncorrected p=0.006, Table II.S5). Overall, the effect of glyphosate on bacterioplankton alpha diversity was relatively weak and not influenced by time (Table II.S5).

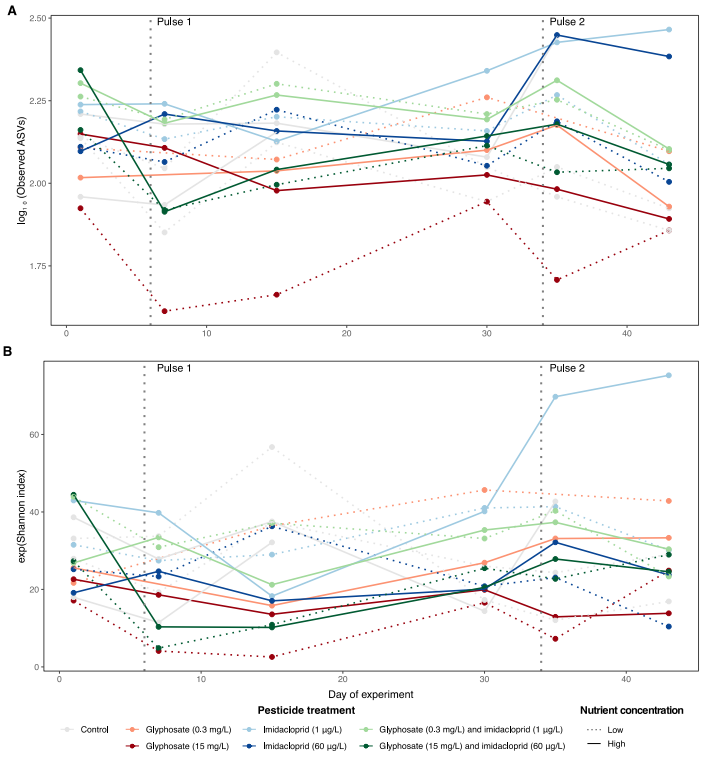


Figure II.4 Bacterioplankton alpha diversity variation across experimental treatments over time, calculated as (A) the (log transformed) observed number of ASVs per sample and as (B) the exponent of Shannon index. Dashed vertical lines indicate days of pesticides pulses application.

We also tracked changes in bacterioplankton community composition, in two ways: with weighted UniFrac distance and JSD, both calculated after normalizing read depth per sample with DESeq2 (or with alternative normalizations described below). We display these changes in community composition using PCoA, with a separate plot for each timepoint of the experiment (Figure II.5). Glyphosate explained a significant proportion of the variation in both metrics of community composition, with R^2 ranging from 0.29 to 0.58, depending on the time following glyphosate application (PERMANOVA, $p<0.007$ for both metrics at all tested timepoints after pesticide pulses, except for weighted UniFrac distance at day 30, Tables S6 and S7). Nutrients and imidacloprid did not significantly affect community composition, alone or in combination with other treatments (Tables S6 and S7). Although nutrients appear to have a slight effect on community composition on day 15 (uncorrected $p=0.027$ for weighted UniFrac and JSD, Table II.S6 and Table II.S7) and on day 30 (uncorrected $p=0.055$ for weighted UniFrac, Table II.S6, and uncorrected $p=0.013$ for JSD, Table II.S7), the effect is not significant after Bonferroni correction, and the explained variance is never as high as it is for glyphosate on the same day ($R^2=0.12$ for both weighted UniFrac and for JSD at both days, Table II.S6 and Table II.S7). We conclude that glyphosate was the dominant driver of compositional changes as it produced a significant and consistent effect on bacterioplankton communities, independent of other stressors, on days 7, 15, 30 and 43 according to JSD, and on days 7, 15 and 43 according to weighted UniFrac distance.

Alternative read depth normalization methods (ASV relative abundance and rarefied data; see Methods) produced qualitatively similar results, showing the predominant effect of glyphosate on community composition (Figure II.S3), with a slight delay in the effect of the first glyphosate pulse compared to the DESeq2 normalization (Figure II.5). The effect of glyphosate on bacterioplankton community composition is detected regardless of the data normalization (Table II.S8), but is more apparent using DESeq2 (compare Figure II.5 to Figure II.S3). This might be because DESeq2 involves a log transformation which reduces the weight of highly abundant community members (McKnight et al., 2019). If less abundant taxa are more responsive to glyphosate, this could explain why this effect is more apparent with DESeq2 normalization.

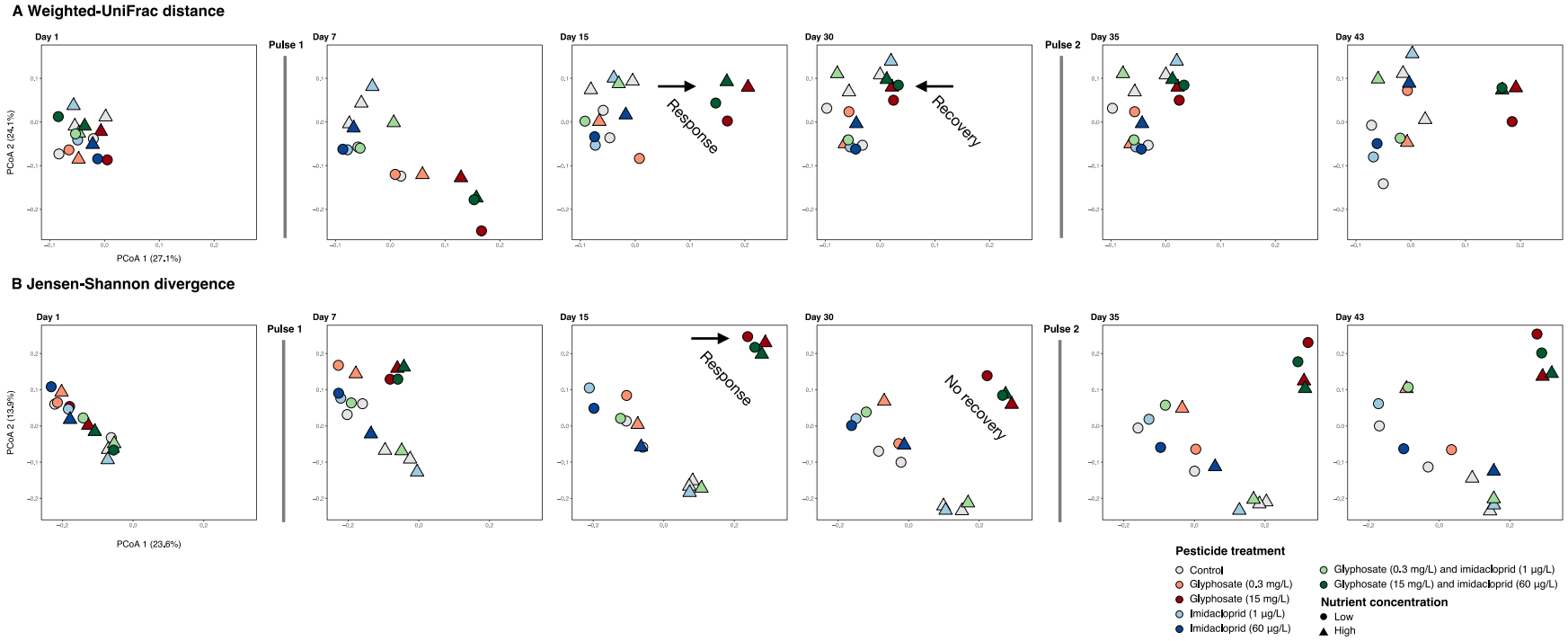


Figure II.5 PCoA ordinations of bacterioplankton community composition in response to experimental treatments, based on (A) weighted UniFrac distance or (B) Jensen-Shannon divergence calculated on ASV estimated absolute abundances after a DESeq2 normalization. Dashed vertical lines indicate days of pesticides pulses application. Each sampling day is plotted in a separate panel to facilitate visualization of treatment effects on community composition, mainly driven by high glyphosate (15 mg/L).

Bacterioplankton communities recover over time at broad phylogenetic scales from the first glyphosate pulse

On day 30 (24 days after the first pesticide pulse and before the second pulse), the bacterioplankton community composition in ponds that had been affected (on day 15) by a high dose of glyphosate (15 mg/L) appeared to recover according to weighted UniFrac (Figure II.5A), but not when using JSD applied to ASVs (Figure II.5B). Using weighted UniFrac, the effect of glyphosate was visibly weaker on day 30 (Figure II.5A) and at the limit of significance after Bonferroni correction (PERMANOVA, R^2=0.29, uncorrected *p*=0.007, Table II.S6), but still significant using JSD (PERMANOVA, R^2=0.34, *p*=0.001, Table II.S7). Viewed together, our series of ordinations show that detection of community recovery depends upon whether phylogenetic information is taken into account. Recovery was apparent when phylogenetic distance among ASVs was calculated (as measured by UniFrac distance, on day 30, control and high-glyphosate communities approach each other, Figure II.5A) but undetected at the ASV level, independent of phylogeny (as measured by JSD, differences between control and high-glyphosate communities keep significant on day 30, Figure II.5B). As such, the community appears to be resilient at a broad phylogenetic level, but not at the finer ASV level, indicating that glyphosate-sensitive ASVs are replaced with phylogenetically-close relatives.

To further assess how resilience varied at different phylogenetic scales, we used PRCs to track community changes at the phylum and ASV levels (Figure II.6). Given that nutrient inputs were not major drivers of community composition (Tables S6 and S7), we built PRCs by combining ponds with the same pesticide treatment, irrespective of nutrient load. This facilitated the visualization of pesticide effects, while capturing the same effects as PRCs considering all experimental treatments separately (compare Figure II.6A and Figure II.S4). We further compared PRCs at different phylogenetic scales, from class to genus level (Figure II.S5). PRCs captured a significant amount of the variation in community responses to pesticide treatments over time (phylum level: F=31.22, class: F=34.28, order: F=26.19, family: F=21.30, genus: F=20.6, ASV: F=10.61, all *p*=0.001; Table II.S9), with greater variation explained at broader taxonomic levels compared to finer levels. The variance explained by the first PRC axis decreased from 47.7% at the

phylum level to 22.1% at the ASV level (Table II.S9). At the broadest taxonomic scale (phylum), communities showed a clear response to high (15 mg/L) but not low (0.3 mg/L) concentrations of glyphosate, followed by a recovery before the second pulse (Figure II.6A). Notably, no recovery was observed at the ASV level (Figure II.6B), consistent with the community composition analysis (Figure II.5). Imidacloprid had no detectable effect at any concentration, whereas the highest concentration of glyphosate caused the greatest effect on bacterioplankton communities. Similar response and recovery patterns were also observed down to the genus level, with progressively weaker recovery at finer taxonomic scales (Figure II.S5). Community composition showed recovery 24 days after the first pulse of glyphosate, but failed to recover after the second pulse (Figure II.5A and Figure II.6A). While this does not exclude the possibility of an eventual recovery, the duration of our experiment (which ended nine days after the second pulse) was likely insufficient to permit subsequent recovery. These results further support that high concentrations of glyphosate led to long-lasting community shifts at the ASV or genus level, whereas community resilience can be achieved at broader phylogenetic scales.

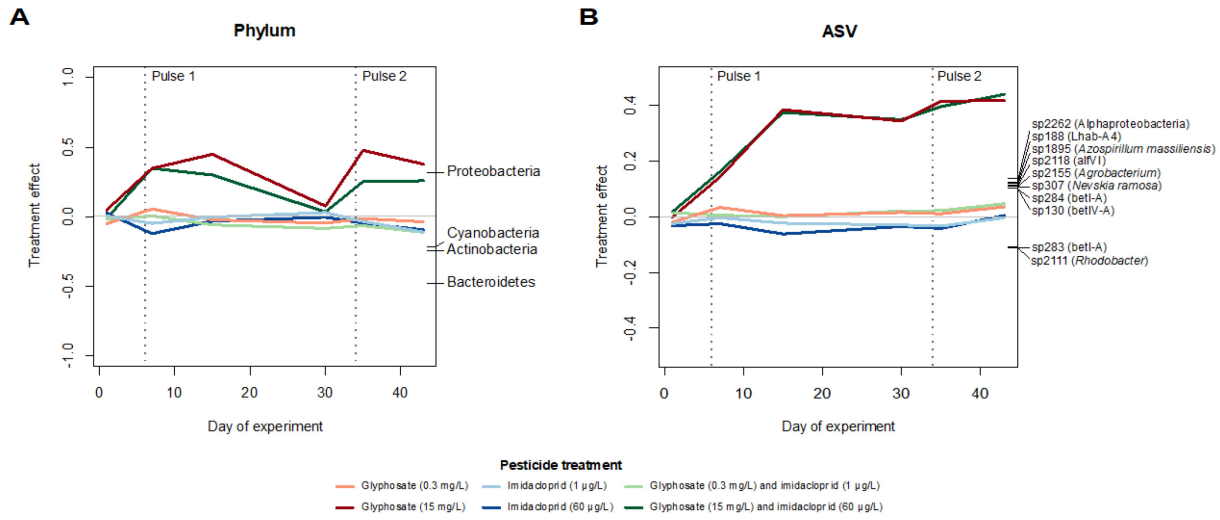


Figure II.6 PRCs showing the effect of pesticide treatments over time relative to control ponds at (A) the phylum level or (B) the ASV level. Dashed vertical lines indicate days of pesticides pulses application. Only phyla with weights >0.2 and ASVs with weight >0.1 are plotted on the right Y axis to facilitate visualization. The finest available annotated level of taxonomic assignment of each ASV is indicated in panel B. Low- and high-nutrient treatments were grouped together for clarity, but follow the same pattern when considered separately (Figure II.S4). See Figure II.S5 for PRCs at other taxonomic levels. These

analyses were based on ASV estimated absolute abundances after a DESeq2 normalization.

Dynamics of the taxa most responsive to treatments

The phylum Proteobacteria was the most positively affected by glyphosate (Figure II.6A; Table II.S10), with relative abundance over 60% in the high glyphosate treatment (15 mg/L) and ~50% or less in other treatments and controls (Table II.S11). Bacteroidetes, Actinobacteria and Cyanobacteria were the most negatively affected phyla (Figure II.6A; Table II.S10, Table II.S11). Of the ten ASVs with the highest absolute taxa weights, all belonged to the phylum Proteobacteria (Figure II.6B, Table II.S12) and, except for sp283 and sp2111, they were all positively affected by glyphosate. An ASV assigned to the genus *Agrobacterium* was among the ASVs that responded most positively to high glyphosate treatment (Figure II.6B; Table II.S12). The GAMM showed that ASVs assigned to the genus *Agrobacterium* increased in EAA over time in ponds receiving high doses of glyphosate (GAMM, factor-smooth interaction between time and high glyphosate treatment: F=19.49, p<0.001, Table II.S13), or receiving both high glyphosate and imidacloprid (GAMM, factor-smooth interaction between time and treatment with high concentrations of both glyphosate and imidacloprid: F=20.66, p<0.001, Table II.S13). A linear time-independent effect of glyphosate was also detected in experimental ponds treated with the highest concentrations of both pesticides together (GAMM, t=7.50, p<0.001, Table II.S13) or glyphosate alone (GAMM, t=6.25, p<0.001, Table II.S13). The modeled *Agrobacterium* abundance (Figure II.S6A) shows a similar 'response followed by recovery' pattern over time as the overall community response at the phylum level (Figure II.6A), suggesting that the positive effect of glyphosate on Proteobacteria may be driven by *Agrobacterium*.

The other two most positively affected genera (*Flavobacterium* and *Azospirillum*, Figure II.S5D) increased in abundance in response to the combination of glyphosate at 15 mg/L and imidacloprid at 60 μg/L (GAMM, factor-smooth interaction between time and treatment with high concentrations of both glyphosate and imidacloprid on *Flavobacterium*: F=17.35, p<0.001, and on *Azospirillum*: F=6.27 p=0.001, Table II.S13) or glyphosate alone at 15 mg/L (GAMM, factor-smooth interaction between time and high

glyphosate treatment on *Flavobacterium*: F=3.63, *p*=0.031, not significant after Bonferroni correction; and on *Azospirillum*: F=5.41, *p*=0.002, Table II.S13), but the effects were not as strong as detected for *Agrobacterium* (Table II.S13). In contrast to the recovery pattern observed in *Agrobacterium* exposed to both the independent and combined highest concentrations of glyphosate (Figure II.S6A), the modeled abundance of *Flavobacterium* (Figure II.S6B) and *Azospirillum* (Figure II.S6C) followed distinct patterns in these two treatments. *Flavobacterium* responded weakly to high doses of both pesticides, mainly after the second pulse, whereas *Azospirillum* recovered partially after responding to the first pulse, but only in ponds treated with the highest concentrations of both pesticides. Despite the overall strong effect of glyphosate on Proteobacteria, these results highlight how different bacterioplankton taxa (including *Agrobacterium* and *Azospirillum* – both Alphaproteobacteria) can show subtly different responses and recovery patterns to pesticides.

Discussion

Context and summary of the experiment

The herbicide glyphosate has been shown to affect aquatic microbial community structure in a variety of natural environments and experimental setups (Berman et al., 2020; Lu et al., 2020; Muturi et al., 2017; Stachowski-Haberkorn et al., 2008). Likewise, the insecticide imidacloprid may disrupt aquatic food webs (Yamamuro et al., 2019), with potential, yet poorly explored consequences for bacterioplankton. The interactive effects of these pesticides on bacterioplankton – and how they might vary depending on fertilizer use and lake trophic status – are relevant because such agrochemical mixtures are common in agriculturally impacted watersheds. Here, we tested how individual and combined gradients of glyphosate and imidacloprid affected bacterioplankton communities in aquatic mesocosms receiving different nutrient inputs. Although they are incomplete representations of natural ecosystems, mesocosm experiments allow us to manipulate and replicate the exposure of complex lake bacterial communities to agricultural chemical pollutants commonly found in freshwaters (Alexander et al., 2016). The current experiment is limited to the response of bacterioplankton communities derived from a pristine lake. Future studies focusing on biofilms and sediments could

complement our results, as many contaminants accumulate in lake sediments and may affect the biofilm community structure (Fernandes et al., 2019; Khadra, Planas, Girard, & Amyot, 2018; Romero et al., 2020).

Glyphosate as a driver of community structure

Our data support the prediction that glyphosate would affect bacterioplankton community structure, which occurred at the highest tested concentration (15 mg/L). Contrary to expectation, no evident interaction between glyphosate and imidacloprid or nutrient load was detected in determining either bacterial density or community structure. High doses of glyphosate resulted in a weak time-independent reduction of bacterioplankton alpha diversity, and a more pronounced change in community composition over time. As hypothesized, glyphosate and nutrient treatments slightly increased bacterial density, suggesting a mild fertilizing effect of glyphosate consistent with it being a potential phosphorus source (Hove-Jensen et al., 2014; Lu et al., 2020). Most bacterioplankton from a pristine source environment (Lake Hertel) are thus able to cope with concentrations of imidacloprid as high as 60 μg/L and of glyphosate as high as 0.3 mg/L, but they may be sensitive to glyphosate concentrations exceeding 15 mg/L. The regulatory criteria intended for eukaryotes (below 60 μg/L for imidacloprid; Table II.S1) were sufficient to preserve bacterioplankton diversity in the experimental conditions at LEAP. On the other hand, the threshold of 15 mg/L for glyphosate deserves further attention from regulatory agencies, as this concentration impacted bacterioplankton composition, which is known to affect lake health and freshwater quality (Kraemer et al., 2020).

Although the highest targeted imidacloprid concentration was not always achieved in all ponds (Figure II.S1), this cannot entirely explain its lack of detectable effect on bacterioplankton. Community composition of ponds receiving measured concentrations of imidacloprid as high as 15 µg/L or more did not deviate from controls, confirming a true lack of effect at least up to that concentration. Alternatively, the absence of a detectable response might be due in part to rapid degradation of imidacloprid in water, which fell below the limit of detection between pulses (Figure II.S1). The absence of a bacterioplankton response is also consistent with the weak or undetectable response of zooplankton biomass to imidacloprid pulses in the same experiment (M.-P. Hébert et al.,

2021). The invertebrate community in the experimental ponds was mainly composed of the zooplanktonic groups Cladocera, Copepoda and Rotifera, and only copepods declined over time after pulse 2, with no resulting effect in total zooplankton biomass (M.-P. Hébert et al., 2021). Overall, these results indicate that the concentrations of imidacloprid applied in this experiment were not sufficient to strongly alter either zooplankton or bacterioplankton biomass or community structure.

Our results suggest that two properties of ecological stability – resistance and resilience – are at play in lake bacterioplankton: functions related to microbial carbon substrate use are resistant to imidacloprid, glyphosate and their interactions in different nutrient backgrounds, while bacterioplankton community composition is resilient following disturbance caused by a glyphosate pulse at 15 mg/L. The recovery of bacterioplankton community composition was only evident when grouping ASVs at higher (more inclusive) taxonomic or phylogenetic levels. Glyphosate thus drove a turnover of bacterioplankton ASVs which, even after the recovery, are different from the ASVs initially found in the undisturbed community.

Proteobacteria are major responders to glyphosate

Glyphosate treatments had a strong positive effect on the phylum Proteobacteria, previously found to be favoured by high concentrations of glyphosate in rhizosphere- (Newman et al., 2016) and phytoplankton-associated communities (C. Wang, Lin, Li, Lin, & Lin, 2017). Multiple species of Proteobacteria can use glyphosate as a source of phosphorus by breaking its C-P bond (Hove-Jensen et al., 2014). We identified *Agrobacterium*, a genus of *Rhizobiaceae* containing species known to degrade glyphosate (Hove-Jensen et al., 2014), as being highly favored in the glyphosate treatment at 15 mg/L. The abundance of ASVs assigned to this genus peaked after each pulse and decreased before the second pulse, coinciding with the community recovery observed 24 days after the first perturbation. The ability to degrade glyphosate may be widespread in the family *Rhizobiaceae* (Liu, McLean, Sookdeo, & Cannon, 1991), and *Agrobacterium* have also been found to encode glyphosate-resistant EPSPS genes (Funke et al., 2006). In fact, this genus was used to create glyphosate-resistant crops, i.e. the so-called ‘Roundup-ready technology’ (Funke et al., 2006). While glyphosate may be a stressor for the microbial community at large (e.g., phytoplankton (Fugère et al.,

2020)), it may be a resource for some members such as *Agrobacterium*, who could potentially detoxify the environment and thus facilitate community recovery after a pulse perturbation. Further genomic and metagenomic analyses of our experimental samples could reveal whether these ecological dynamics are underlain by evolutionary adaptation, and whether community resistance and resilience can be explained by the initial presence of resistant bacteria in the community, or to *de novo* mutations or gene transfer events.

Glyphosate could have driven changes in the bacterial community via direct mechanisms (e.g. by affecting species with a sensitive EPSPS, its target enzyme) or indirect mechanisms (e.g. effects on other trophic levels that cascaded down to bacteria via predation or other interactions). In a previous study of the same experiment described here that focused on the responses of eukaryotic phytoplankton, we found that glyphosate treatment reduced the diversity of phytoplankton, but did not significantly change phytoplankton community composition (Fugère et al., 2020). Although a reduced phytoplankton diversity could indirectly affect bacterioplankton community composition, a direct effect of glyphosate on bacteria seems more plausible as the taxa favored by the treatment (mainly Proteobacteria) have been previously shown to be directly affected in a similar way (Janßen et al., 2019; C. Wang et al., 2017). Indeed, bacterial degradation of glyphosate likely released bioavailable phosphorus, stimulating phytoplankton growth (Fugère et al., 2020). Further studies will be needed to disentangle how the effects of pesticides cascade through food webs, and how trophic structure influences their effects.

Functional redundancy in carbon utilization potential

Despite the marked changes in taxonomic composition driven by glyphosate, microbial communities did not change their carbon substrate use throughout the experiment, providing evidence for functional redundancy in carbon utilization potential. This was an expected result, as broad-scale ecosystem functions such as respiration and dissolved organic carbon consumption are weakly coupled with species composition (Girvan, Campbell, Killham, Prosser, & Glover, 2005; Langenheder, Lindström, & Tranvik, 2006; Peter et al., 2011), allowing these functions to remain unaffected by fluctuations in microbial community composition (Louca et al., 2018). While less diverse communities (in terms of species richness) may lack functional redundancy, more diverse communities are expected to encode more redundant functions (Konopka, 2009). We can thus surmise

that the freshwater bacterioplankton communities studied here were sufficiently diverse to be functionally redundant for carbon utilization in the face of disturbance. The weak and time-independent effect of high concentrations of glyphosate on alpha diversity was insufficient to alter community carbon substrate use. However, our experiment was conducted with communities originating from a pristine lake in a nature reserve, and this result might not be generalized to freshwaters historically impacted by other forms of anthropogenic stress. For example, land use intensity is negatively correlated with bacterioplankton richness in lakes across Eastern Canada (Kraemer et al., 2020). It remains to be seen whether such impacted lakes are less functionally redundant, and thus possibly more susceptible to impaired ecosystem functioning. Lastly, although bacterioplankton respiration accounts for a large fraction of organic carbon processing in freshwaters (Berggren, Lapierre, & del Giorgio, 2012), the carbon substrate use we measured could also be due in part to fungal activity which could be compensating or masking changes in bacterioplankton activity. There was no macroscopically observable fungal growth in the plates, yet microscopic fungi likely contributed a fraction of the inoculum used to initiate the plates.

The phylogenetic depth of glyphosate resistance: Methodological considerations

The inference of bacterioplankton ASVs in this study allowed a relatively fine-scale taxonomic resolution of community changes in response to a pulse perturbation of glyphosate. Notably, the recovery of bacterioplankton composition was detectable at broader taxonomic units (e.g. phylum in particular) but not at the ASV level. This implies that the taxonomic resolution of traits under selection during recovery from a glyphosate pulse is relatively coarse (Martiny, Jones, Lennon, & Martiny, 2015). This result could also be explained if ASVs are too fine-scale as a measure of diversity, and mostly reflect sequencing or base calling errors rather than true biological diversity. We deem this unlikely, first because the ASV inference algorithm includes a model-based approach to correct for amplicon sequencing errors (Callahan et al., 2016), and second because ASV detection methods are usually more accurate than OTU-clustering methods based on sequence similarity thresholds of usually 97% (Caruso, Song, Asquith, & Karstens, 2019). For example, we only found 7 to 10 false-positive ASVs (Methods), but dozens to hundreds of false positive are detected by even state-of-the-art (distribution-based)

sequence clustering-based methods to identify operational taxonomic units, when applied to the same or similar mock communities as used here (Tromas et al., 2017). Although we cannot exclude the impact of possible false ASVs on our results, we expect them to be relatively minimal and evenly distributed across all timepoints (Callahan, McMurdie, & Holmes, 2017). In other words, there is no reason to believe that sequencing errors should be non-randomly distributed over time or across experimental treatments. Moreover, PRC analyses show a steady decline from the phylum level to the genus level in both the response to, and recovery from, high concentrations of glyphosate. Therefore, even without considering the ASV level, there is still a discernible pattern of greater community resilience at broader taxonomic scales. This suggests that the traits (and underlying genes) required for survival or growth in the presence of glyphosate are relatively deeply conserved. Higher-resolution genomic or metagenomic analyses could be used to confirm this result, and pinpoint the genes involved in resistance.

Ecosystem resistance, resilience and stability

Our study provides evidence of ecosystem stability in terms of carbon substrate use maintained by microbial communities when faced by a perturbation by two of the most commonly used pesticides in the world, separately or in combination. We also showed resistance to a wide gradient of imidacloprid contamination, and resilience to high doses of glyphosate in bacterioplankton communities that have no known history of contact with the herbicide. Finally, whether a stressed community is considered resilient depends on the phylogenetic depth of the traits required to deal with the stress (Martiny et al., 2015). Our results provide an example of how resilience to stressors can be a feature of deeper phylogenetic groups, but not finer-scale groupings (ASVs), which could be involved in adaptation to other stressors or niches.

Acknowledgements: We are grateful to D. Maneli, C. Normandin, A. Arkilanian and T. Jagadeesh for their assistance in the field, to J. Marleau, C. Girard, O.M. Pérez-Carrascal and N. Tromas for their assistance during laboratory analyses, to K. Velghe for nutrient analyses and to M.A.P. Castro for developing the LC–MS method for pesticides quantification. We thank the three anonymous reviewers for comments on a previous version of this manuscript.

Data accessibility: Sequence data corresponding to raw 16S rRNA reads and metadata, as well as carbon substrate utilization dataset based on Biolog EcoPlates assessment, are available on https://figshare.com/projects/MEC-LEAP/78297. Sequences of 16S rRNA reads are also available at NCBI SRA (BioProject ID PRJNA664121).

Funding information: This study was supported by a Canada Research Chair and NSERC Discovery Grant to B.J.S. N.B.C. was funded by FRQNT and NSERC-CREATE/GRIL fellowships. V.F. was supported by an NSERC postdoctoral fellowship. M-P.H. was funded by NSERC and NSERC-CREATE/GRIL. C.C.Y.X. was funded by a Vanier Canada Graduate Scholarship. R.D.H.B. was supported by a Canada Research Chair. LEAP was built and operated with funds from a CFI Leaders Opportunity Fund, NSERC Discovery Grant and the Liber Ero Chair to A.G.

Author contributions: N.B.C., V.F., M.-P.H., R.D.H.B., B.E.B., G.B., G.F.F., B.J.S. and A.G. designed the study. N.B.C., V.F. and M.-P.H. collected the data. N.B.C., C.C.Y.X. and V. Y. contributed to the development of laboratory methods. N.B.C. and V.F. analysed data. N.B.C. made the figures and drafted the manuscript. N.B.C. and B.J.S. wrote the first manuscript draft and all authors contributed significantly to data interpretation and commented on manuscript drafts.

Supplementary information

Supplementary methods

Illumina library preparation and sequencing

We performed a two-step PCR amplification of the V4 region of the 16S rRNA gene to first amplify the region (step 1) and then attach barcodes and Illumina adapters to the amplicon (step 2) following protocol described in Preheim et al. (2013). All PCR reactions were performed in Mastercycler nexus thermocyclers (Eppendorf Corporate, Mississauga, Canada).

Step 1 PCR was performed with 0.5 unit of Phusion DNA polymerase and 1X Phusion High Fidelity Buffer (ThermoFisher, Waltham, MA, USA), 200 µl of dNTPs (ThermoFisher, Waltham, MA, USA), 0.36 µM of each primer (U515_F and E786_R) and 20 ng of extracted environmental DNA in reactions of 25 µl. PCR conditions were: initial denaturation for 30 seconds at 98 ˚C, followed by 22 cycles of denaturation at 98 ˚C for 20 seconds, annealing at 54 ˚C for 35 seconds and extension at 72 ˚C for 30 seconds, the cycles were followed by final elongation at 72 ˚C for 60 seconds. Four reactions of 25 µl were performed per sample, pooled and cleaned with the Zymo research DNA purification kit (Zymo Research, Irvine, USA) according to manufacturer's protocol.

Step 2 PCR was performed similarly to step 1, but with 4 µl of the purified step 1 PCR product and 0.36 µM of PE-III-PCR-F and 0.36 µM of barcoded reverse primers PE-III-PCR-XXXX (see exact barcode sequence of each sample in Table II.S3). PCR conditions were: initial denaturation for 30 seconds at 98˚C, followed by seven cycles of 98˚C for 30 seconds, 83˚C for 30 seconds and 72˚C for 30 seconds. Reactions were performed in duplicates, then pooled and purified using Agencourt AMPure XP beads (Beckman Coulter Life Sciences, Indianopolis, IN, USA) following manufacturer's instructions. Fragment size was confirmed via electrophoresis in agarose gels and quantified on a NanoDrop microvolume spectrophotometer (ThermoFisher, Waltham, MA, USA).

Libraries were then pooled at equimolar ratio, denatured and sequenced, using the MiSeq reagent Kit V2 with 500 cycles (Illumina, San Diego, CA, USA) yielding two 250 bp paired-end reads. The 96 samples were split between two different sequencing runs, each of them including PCR negative controls, extraction blanks, and a mock community

DNA sample containing identified 16S rRNA clone libraries. PCR negative controls consisted of ultrapure DNase/RNase-free distilled water (ThermoFisher, Waltham, MA, USA), and DNA extraction blanks from clean filters. In one of the runs we used a custom mock community composed of 16S rRNA clone libraries from freshwater lake samples (Preheim et al. 2013) and in the other we used the American Type Culture Collection MSA-1002 mock community (ATCC, Manassas, VA, USA).

Supplementary figures

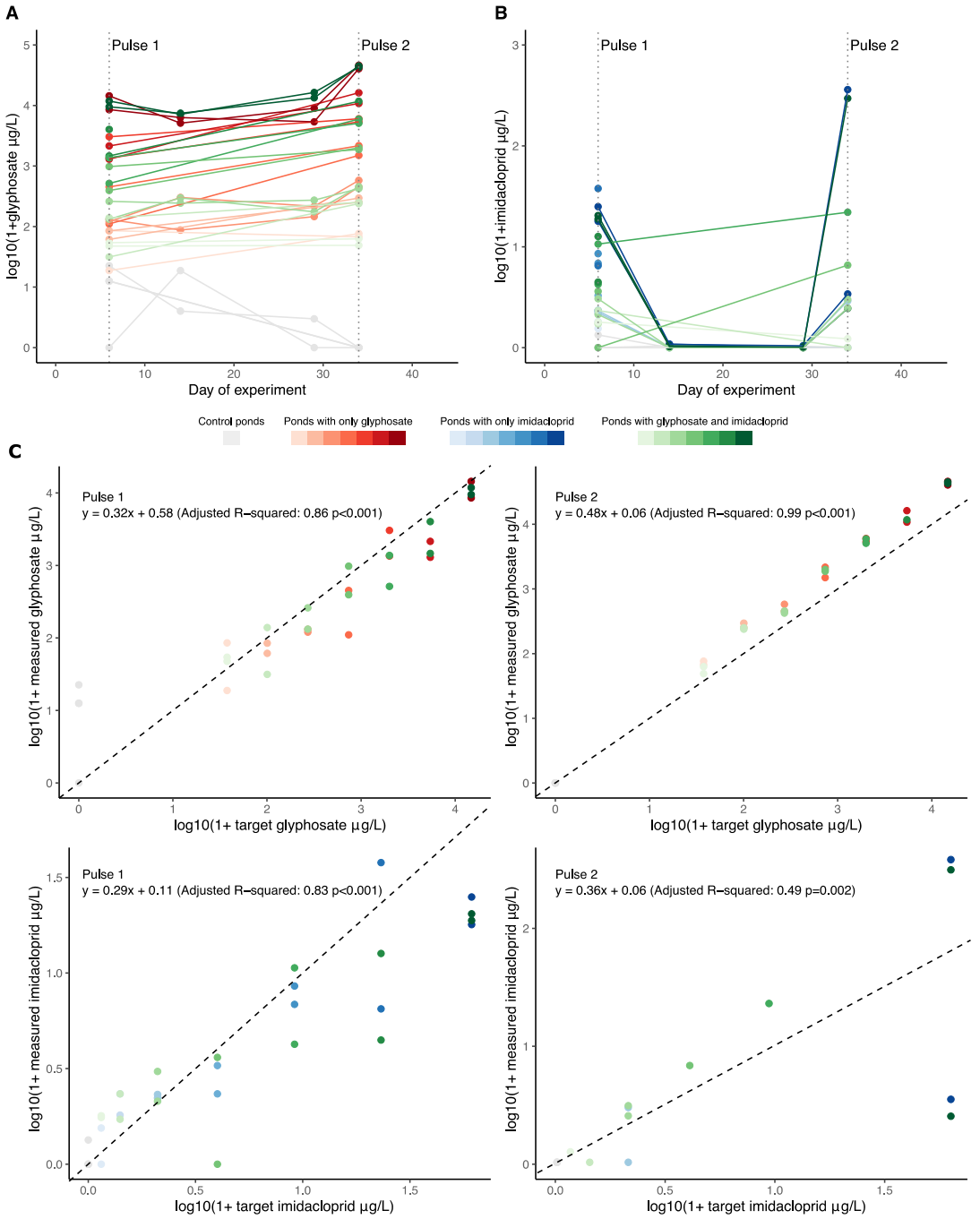


Figure II.S1 Experimental gradient established for (A) glyphosate and (B) imidacloprid concentrations between two application pulses (at days 6 and 34) and (C) the correlation between target and measured concentrations at each pulse. The top row of figure C shows results for glyphosate, and the bottom two rows for imidacloprid, after pulse 1 (left column) and pulse 2 (right column) respectively.

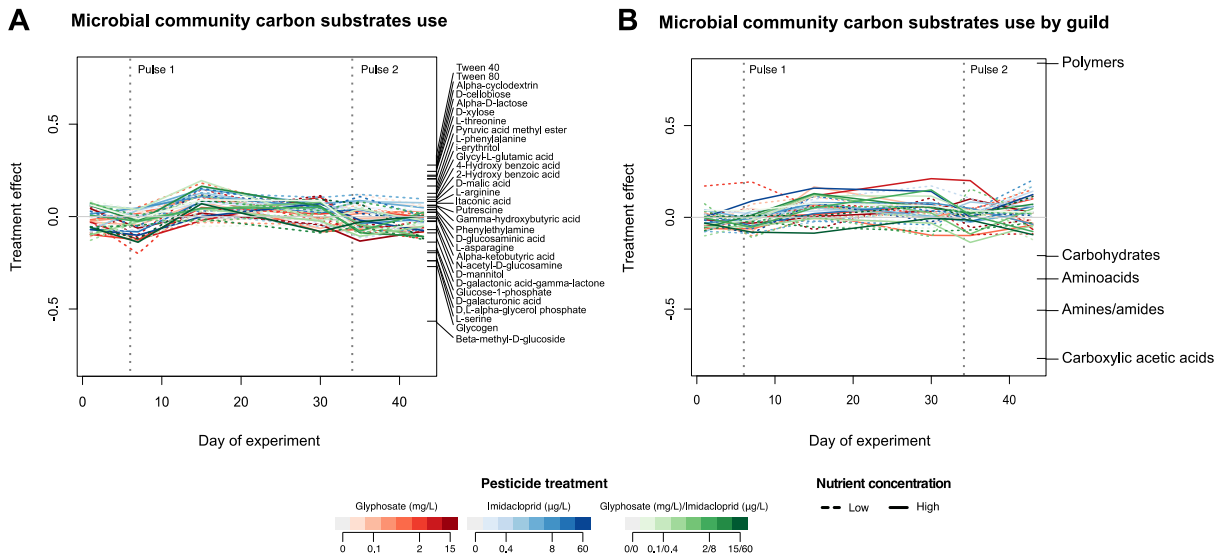


Figure II. S2 PRC plots show no effect of experimental treatments on community metabolic profiles when considering (A) the 31 different compounds individually (F=32.6 p=0.69) or (B) grouped according to functional guilds (F=79.2 p=0.86). The PRC axis shown in A explains 13.4% of total variance and in B 43.1%.

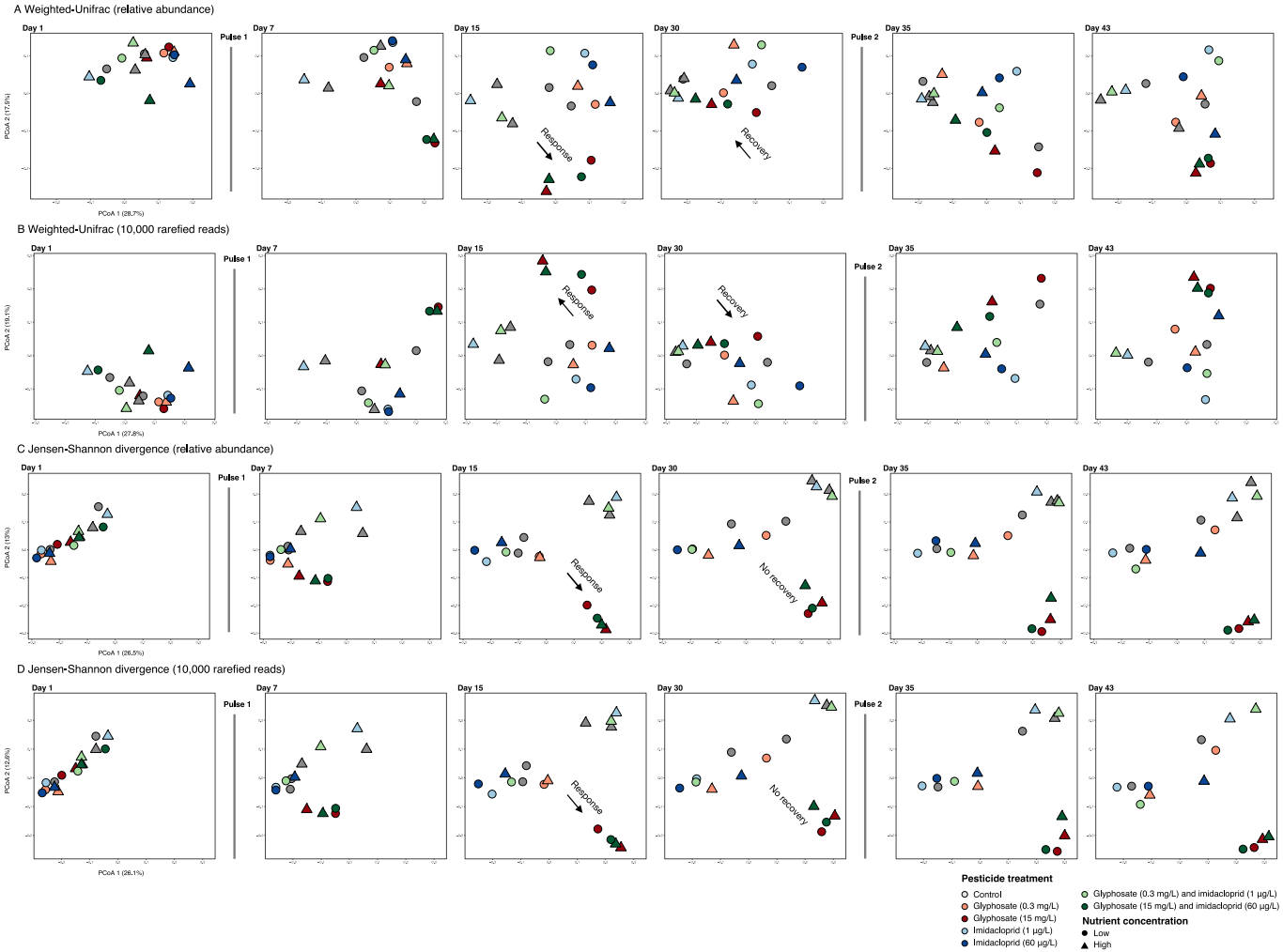


Figure II. S3 PCoA ordinations based on (A, B) weighted UniFrac distance or (B, D) Jensen-Shannon divergence exploring different normalization approaches: (A, C) calculation of reads relative abundance and (B, D) rarefying to a threshold of 10,000 reads per sample. Each sampling day is plotted separately to facilitate visualization of treatment effects on community composition.

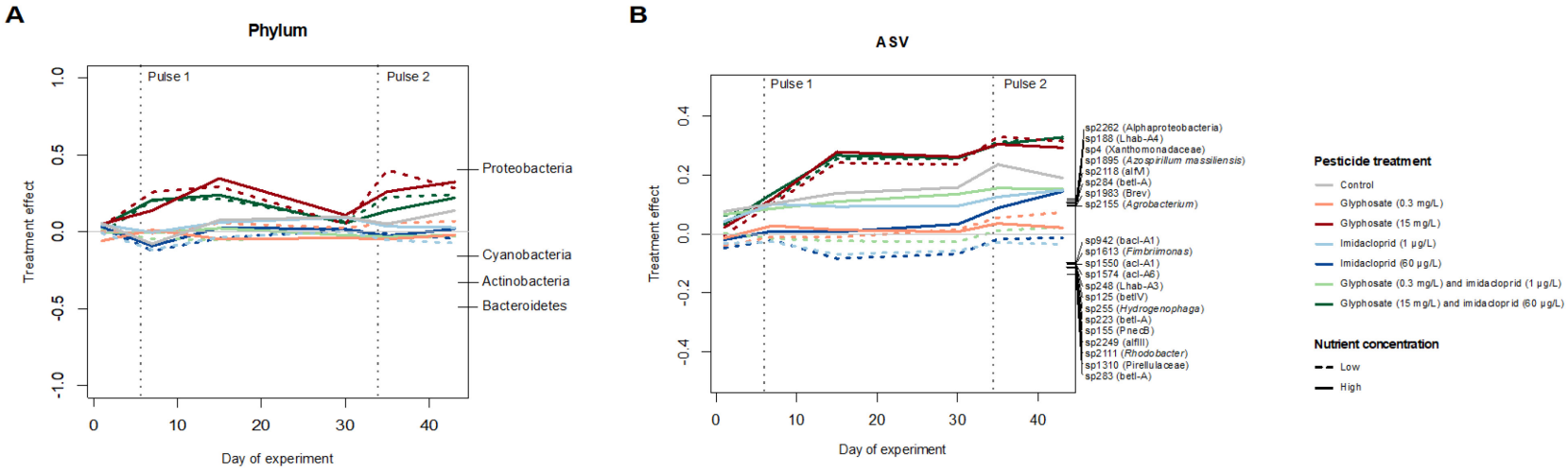


Figure II. S4 PRCs show how high glyphosate treatments affected community composition at (A) phylum and (B) ASV levels. Low- and high-nutrient treatments show the same pattern, and, for this reason, they were grouped in Figure II.4, to facilitate data visualization. The finest level of taxonomic assignment based on FreshTrain and GreenGenes database is shown following ASV names in panel B. Only taxa with weights higher than 0.2 are shown in A and higher than 0.095 are shown in B.

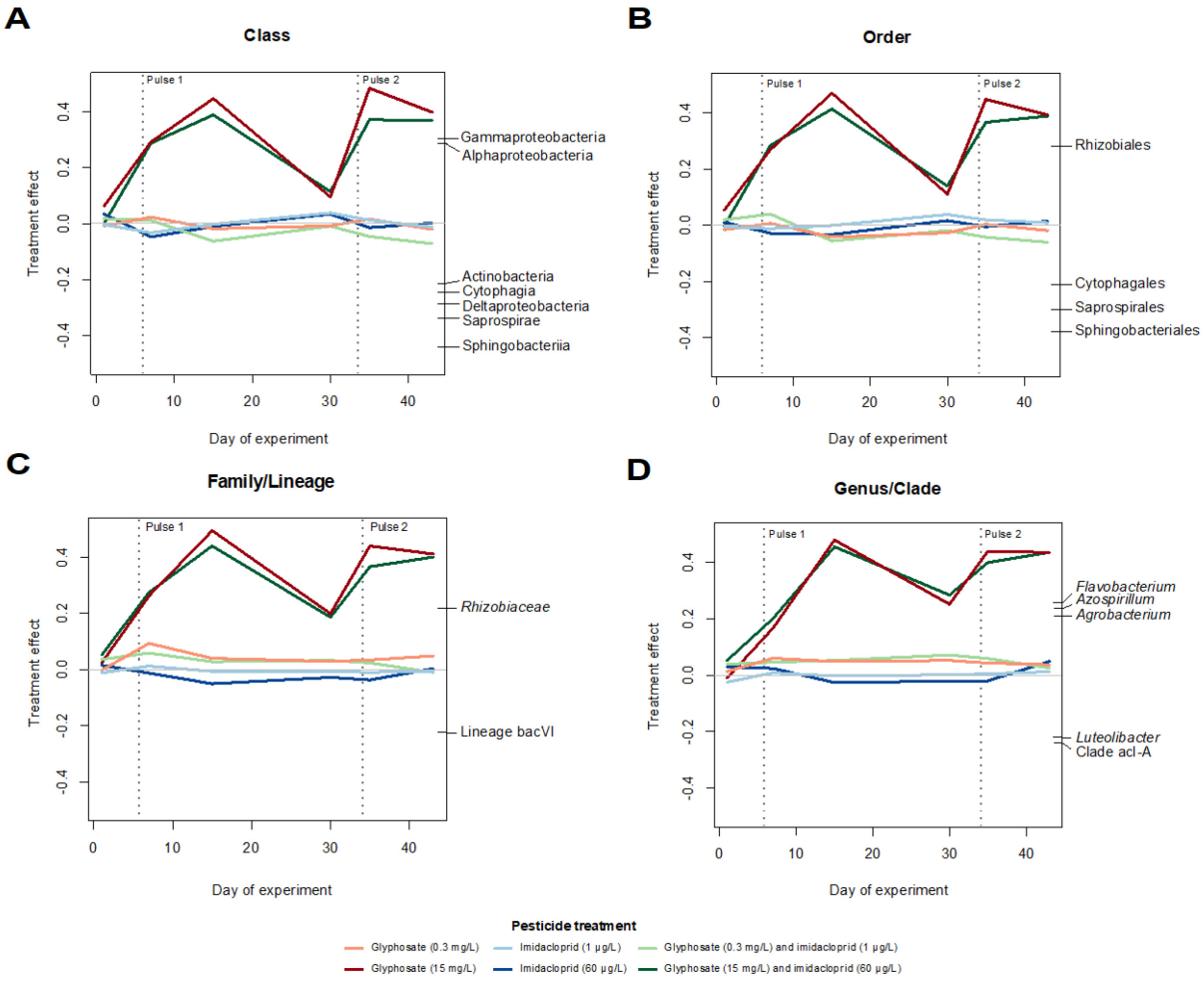


Figure II. S5 PRCs show how high glyphosate treatments (15 mg/L) affected community composition at different taxonomic levels: (A) class, (B) order, (C) family/lineage, (D) genus/clade. Taxonomic assignment based on FreshTrain and GreenGenes databases. Low and high nutrient treatments were grouped as they follow the same pattern. Only taxa with weights higher than 0.2 are shown.

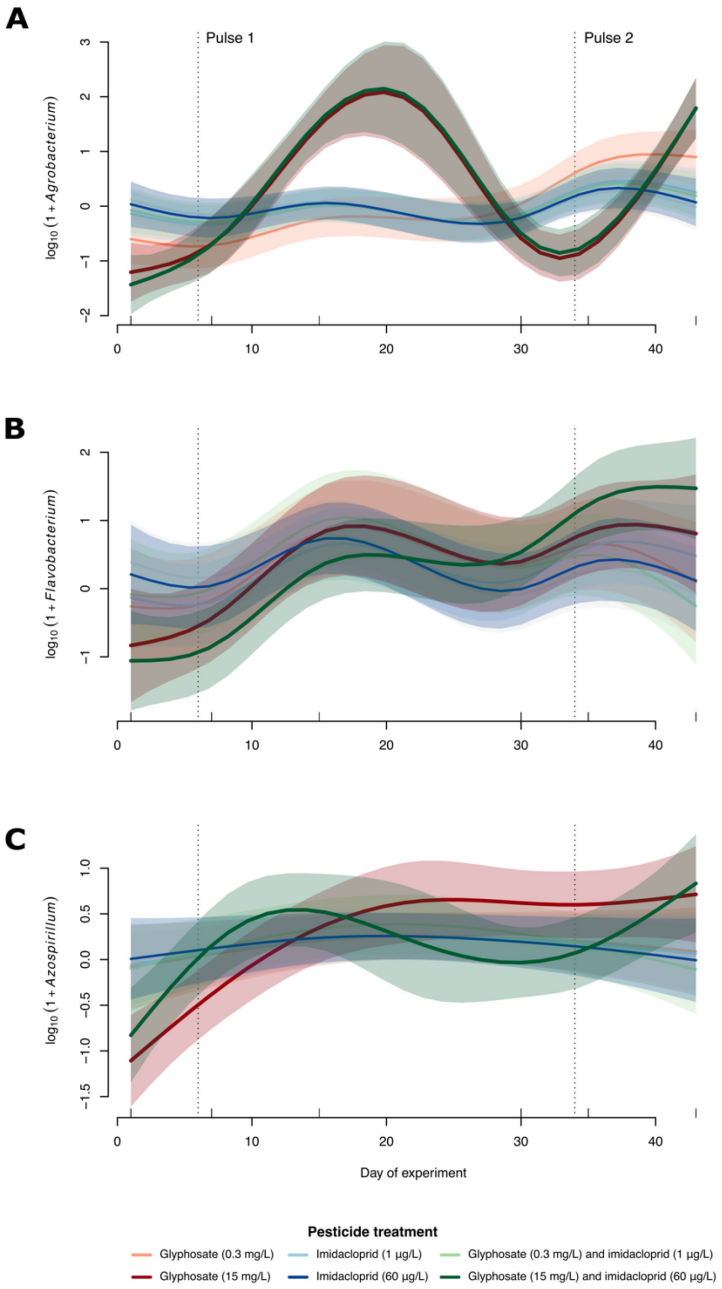


Figure II. S6 Summed effects of GAMMs on abundance of three genera most positively affected by the glyphosate treatments: (A) *Agrobacterium*, (B) *Flavobacterium* and (C) *Azospirillum*. Shades indicate a confidence interval of 95%. Abundance of each genus is the estimated absolute abundance of all ASVs assigned to *Agrobacterium*, Flavobacterium or *Azospirillum* after normalization by rarefying each sample to 10,000 reads without replacement.

Supplementary tables

Table II.S1 Regulatory acceptable concentrations (RACs) of glyphosate and imidacloprid residues in freshwater according to regulatory agencies in Canada (CCME, Canadian Council of Ministers of the Environment), Europe (EFSA, European Food Safety Agency) and in the USA (EPA, Environmental Protection Agency). Chronic (long-term) and acute (short-term) exposure RACs are specified when available.

Regulatory agency	Glyphosate	Imidacloprid
CCME (Canada)	800 µg/L chronic 27,000 µg/L acute	0.23 µg/L interim
EFSA (Europe)	Risk to aquatic organisms considered low Selected data from toxicological studies: - 12,500 µg/L chronic toxicity to *Daphnia magna* - 40,000 µg/L acute toxicity to *D. magna* - 8,500 µg/L acute toxicity to *Aphanizomenon flosaquae*	Tier 1† - 0.209 µg/L chronic - 0.341 µg/L acute Tier-2B† - 0.009 µg/L chronic - 0.098 µg/L acute
EPA (USA) ‡	26,600 µg/L chronic 49,900 µg/L acute	0.01 µg/L chronic 0.385 µg/L acute

† Tier 1 is indicated as not appropriate for risk assessment and Tier-2B to be used provisional risk for assessment

‡ EPA's Office of Pesticide Programs (OPP) for aquatic invertebrates

Table II.S2 Carbon substrates present in Biolog EcoPlates and their respective grouping (guild)

Substrate	Group (guild)
Phenylethylamine	Amines/amides
Putrescine	Amines/amides
L-arginine	Amino acids
L-asparagine	Amino acids
L-phenylalanine	Amino acids
L-serine	Amino acids
L-threonine	Amino acids
Glycyl-L-glutamic acid	Amino acids
Pyruvic acid methyl ester	Carbohydrates
D-cellobiose	Carbohydrates
Alpha-D-lactose	Carbohydrates
Beta-methyl-D-glucoside	Carbohydrates
D-xylose	Carbohydrates
i-erythritol	Carbohydrates
D-mannitol	Carbohydrates
N-acetyl-D-glucosamine	Carbohydrates
Glucose-1-phosphate	Carbohydrates
D,L-alpha-glycerol phosphate	Carbohydrates
D-glucosaminic acid	Carboxylic acetic acids
D-galactonic acid-gamma-lactone	Carboxylic acetic acids
D-galacturonic acid	Carboxylic acetic acids
2-Hydroxy benzoic acid	Carboxylic acetic acids
4-Hydroxy benzoic acid	Carboxylic acetic acids
Gamma-hydroxybutyric acid	Carboxylic acetic acids
Itaconic acid	Carboxylic acetic acids
Alpha-ketobutyric acid	Carboxylic acetic acids
D-malic acid	Carboxylic acetic acids
Tween 40	Polymers
Tween 80	Polymers
Alpha-cyclodextrin	Polymers
Glycogen	Polymers

Table II.S3 Barcode sequences of the reverse primer used in step 2 PCR, and total read counts obtained after sample demultiplexing. The number of non-chimeric reads obtained after filtering, denoising, merging paired ends and removing chimeras with DADA2, is also shown.

Sample	Barcode sequence	Total reads	Non-chimeric reads
C1.17.aug	AACCCGTT	55762	50337
C1.23.aug	CAGCGGCA	17714	15227
C1.28.sep	GTTCGCAG	14880	13926
C1.31.aug	CCGACAAA	51823	44758
C1.15.sep	TAAGGGAG	12294	11295
C1.20.sep	TTGTGGCG	23611	22246
C4.17.aug	GACATCAT	56234	50201
C4.23.aug	GAGTTTGA	7958	7118
C4.31.aug	AACAGTAT	15241	11601
C4.15.sep	ATCGCACC	39833	36356
C4.20.sep	CTAGAATC	12371	7857
C4.28.sep	CCTTTGAT	23931	21410
C8.17.aug	ATAGGTGG	18402	16033
C8.23.aug	CAACTTCA	29342	27344
C8.31.aug	GTAGTCGA	21294	19270
C8.15.sep	TCCCGATG	19440	18054
C8.20.sep	GGGCGAAA	13574	11480
C8.28.sep	GGTGTACC	16869	14077
D1.17.aug	CGTCCCAC	37000	33163
D1.23.aug	CTGTTAGT	36130	33232
D1.31.aug	CACTCACT	32340	29168
D1.15.sep	ACCTCCCA	43478	40627
D1.20.sep	AATACAGG	66220	60120
D1.28.sep	AGTCACCC	9413	8460
D4.17.aug	TACGATAC	24498	21285
D4.23.aug	AGGCTTCA	10161	9333
D4.31.aug	GTGCTGAT	39070	34550
D4.15.sep	ACCATACT	38941	36695
D4.20.sep	AAATTGGA	44426	40952
D4.28.sep	TTCCAGAT	12791	11531
D8.17.aug	GCCTGTTC	35891	32901
D8.23.aug	TCGGCTCG	85995	80978
D8.31.aug	AGAGAGGC	30980	24001
D8.15.sep	GCAATGGA	29363	26644

D8.20.sep	TGACTTAG	34660	31171
D8.28.sep	GGAGGCTG	21580	16647
E1.17.aug	CCGCACCG	27000	24008
E1.23.aug	ATGCCAGC	15379	13891
E1.31.aug	TCGAACAC	68782	63125
E1.15.sep	CGACATTC	26909	20436
E1.20.sep	CATCGCTA	41282	34865
E1.28.sep	ACTAAGAT	18204	15298
E4.17.aug	TCAAAGCT	70496	61775
E4.23.aug	CAGCGGCA	49463	44749
E4.31.aug	CCGACAAA	55130	49059
E4.15.sep	TAAGGGAG	50914	46906
E4.20.sep	TTGTGGCG	49999	44315
E4.28.sep	GTTCGCAG	47427	42091
E8.17.aug	TGGGACCT	61084	52159
E8.23.aug	GAGTTTGA	28586	25272
E8.31.aug	AACAGTAT	54729	48638
E8.15.sep	ATCGCACC	45603	39995
E8.20.sep	CTAGAATC	36482	32974
E8.28.sep	CCTTTGAT	65278	59393
H1.17.aug	TGTTTCCC	12056	10311
H1.23.aug	GGTAATGA	15470	13598
H1.31.aug	GTACGTTG	48086	44387
H1.15.sep	ACGGGCTG	11257	9873
H1.20.sep	ATGAAGTA	12745	11221
H1.28.sep	ACACCTCG	10727	9949
H4.17.aug	ATAGGTGG	38081	35342
H4.23.aug	CAACTTCA	53637	49787
H4.31.aug	GTAGTCGA	40116	36626
H4.15.sep	TCCCGATG	65142	60387
H4.20.sep	GGGCGAAA	29985	27212
H4.28.sep	GGTGTACC	38098	35100
H8.17.aug	CGTCCCAC	71142	62412
H8.23.aug	CTGTTAGT	64022	57820
H8.31.aug	CACTCACT	76152	66286
H8.15.sep	ACCTCCCA	72837	66412
H8.20.sep	GAGCACAG	84816	77740
H8.28.sep	CGAATATT	65034	58318
J4.17.aug	TACGATAC	75103	67734

J4.23.aug	AGGCTTCA	54008	48385
J4.31.aug	GTGCTGAT	54083	50492
J4.15.sep	ACCATACT	62067	57735
J4.20.sep	AAATTGGA	65309	59721
J4.28.sep	TTCCAGAT	30185	27734
J8.17.aug	GCCTGTTC	72714	63613
J8.23.aug	AGCTGACG	58664	50262
J8.31.aug	AGAGAGGC	55708	48142
J8.15.sep	GCAATGGA	58161	52962
J8.20.sep	TGACTTAG	64244	56997
J8.28.sep	GGAGGCTG	43352	37922
K4.17.aug	CCGCACCG	80953	74716
K4.23.aug	ATGCCAGC	81279	74020
K4.31.aug	TCGAACAC	96033	88819
K4.15.sep	CGACATTC	44684	41105
K4.20.sep	CATCGCTA	48869	44067
K4.28.sep	ACTAAGAT	64643	59626
K8.17.aug	TGTTTCCC	80622	73294
K8.23.aug	GGTAATGA	76790	69163
K8.31.aug	GAAACTGG	30648	27342
K8.15.sep	ACGGGCTG	45460	42448
K8.20.sep	ATGAAGTA	42178	37273
K8.28.sep	ACACCTCG	51882	45344

Table II.S4 Summarized results of the generalized additive mixed models (GAMMs) for bacterial density and number of carbon substrate used as a response variables. A Gaussian residual distribution was used for both models. For each response variable we report the sample size (n), adjusted R2, the predictors used in the model, the parameter estimate and respective standard error (SE) of parametric effects or the effective degrees of freedom (EDF) of smooth terms, the corresponding test statistics (t value for parametric and F for smooth terms) and the p-value. Smooths terms are described as mgcv syntax: 's()' functions are thin plate regression splines and 'ti()' tensor product interactions, pond represents the random variable of the mixed model and 'bs='fs'' specifies the underline base function as a random smooth. Following a Bonferroni multiple testing correction for 9 tests, we only considered significant variables with unadjusted p-value <0.005 (shown in bold).

Response variable	n	Adjusted R^2	Predictors (significant effects in bold)	Estimate (SE) or EDF	Statistic	*p*-value
log_{10}(Bacterial density)	288	0.734	**nutrient** **ti(day)** **ti(day, glyphosate)** ti(day, glyphosate, by = nutrient) ti(day, imidacloprid) ti(day, imidacloprid, by = nutrient) ti(day, glyphosate, imidacloprid) ti(day, glyphosate, imidacloprid, by = nutrient) **s(day, Pond, bs='fs')**	**0.08** **(SE=0.02)** **4.6** **4.1** 1.0 2.4 1.0 0.0 3.1 **70.5**	**4.1** **17.5** **6.6** 5.7 0.2 0.5 0.0 2.1 **2.2**	**< 0.001** **< 0.001** **< 0.001** 0.018 0.916 0.490 0.183 0.260 **< 0.001**
Number of carbon substrates used (EcoPlates)	288	0.309	nutrient **ti(day)** ti(day, glyphosate) ti(day, glyphosate, by = nutrient) ti(day, imidacloprid) ti(day, imidacloprid, by = nutrient) ti(day, glyphosate, imidacloprid) ti(day, glyphosate, imidacloprid, by = nutrient) **s(day, Pond, bs='fs')**	0.88 (SE=0.37) **4.5** 1.0 3.1 1.0 6.0 0.0 1.6 **33.4**	2.4 **6.0** 1.0 1.8 1.4 1.3 0.0 1.1 **0.4**	0.019 **< 0.001** 0.320 0.136 0.242 0.244 0.266 0.449 **< 0.001**

Table II.S5 Summarized results of the generalized additive mixed models (GAMMs) for alpha diversity: observed ASV and exponential Shannon. Gaussian residual distributions were used in all models. For each response variable we report the sample size (n), adjusted R2, the predictors and factors used in the model, the parameter estimate and respective standard error (SE) of parametric effects or the effective degrees of freedom (EDF) of smooth terms, the corresponding test statistics (t value for parametric and F for smooth terms) and the p-value. Smooths terms are described as mgcv syntax: 's()' functions are thin plate regression splines and 'ti()' tensor product interactions, pond represents the random variable of the mixed model and 'bs='fs'' specifies the underline base function as a random smooth. Following a Bonferroni multiple testing correction for 16 tests, we only considered significant variables with p-value <0.003, shown in bold.

Response variable	n	Adjusted R^2	Predictors	Factors† of parametric and smooth terms	Estimate (SE) or EDF	Statistic	*p*-value
log_{10}(Observed ASVs)	89	0.69	treatment	high_both	-0.02 (0.07)	-0.27	0.786
				high_glypho	-0.20 (0.07)	-2.89	0.006
				high_imid	0.07 (0.07)	1.06	0.296
				low_both	0.13 (0.07)	1.88	0.065
				low_glypho	-0.02 (0.07)	-0.26	0.795
				low_imid	0.14 (0.07)	2.03	0.048
			nutrient	high	-0.09 (0.04)	2.21	0.032
			ti(day)		4.7	3.18	0.021
			ti(day, by=treatment)	low_glypho	2.4	0.48	0.589
				low_imid	1.0	0.51	0.509
				low_both	1.0	0.21	0.652
				high_glypho	2.0	0.94	0.539
				high_imid	1.0	0.51	0.478
				high_both	3.6	2.49	0.110
			ti(day, by=nutrient)	high	11.2	1.34	0.372
			s(day, Pond, bs='fs')		**1.4**	**1.17**	**<0.001***

exp(Shannon index)	89	0.66	**treatment**	high_both	-7.3 (4.3)	-1.70	0.096
				high_glypho	**-14.9 (4.2)**	**-3.51**	**0.001***
				high_imid	-5.1 (4.2)	-1.22	0.228
				low_both	4.1 (4.3)	0.95	0.346
				low_glypho	1.2 (4.4)	0.28	0.780
				low_imid	11.4 (4.3)	2.64	0.011
			nutrient	high	0.9 (2.5)	0.37	0.715
			ti(day)		4.0	1.62	0.262
			ti(day, by=treatment)	low_glypho	1.0	1.95	0.169
				low_imid	3.7	3.05	0.018
				low_both	2.5	0.46	0.706
				high_glypho	1.7	0.77	0.526
				high_imid	1.0	0.00	0.993
				high_both	3.3	4.36	0.012
			ti(day, by=nutrient)	high	2.2	2.60	0.116
			s(day, Pond, bs='fs')		**10.2**	**1.00**	**0.001***

Table II.S6 PERMANOVA for different explanatory variables (and their interaction) in models with the weighted UniFrac distances among communities as the response. The same model was tested at five different time points and an asterisk indicates p-values that are significant after a Bonferroni correction for 7 hypothesis tests (i.e. p-values <0.007 are considered significant, shown in bold). A PERMDISP was performed to confirm homogeneity of groups dispersions and significant p-values (<0.05) point out to predictors whose significance in the PERMANOVA output should be carefully analysed as they may be a result of within-group variation rather than among-group variation. df=degrees of freedom

			PERMANOVA (999 permutations)			PERMDISP (999 permutations)	
	Predictor	df	F	R^2	*p*-value	F	*p*-value
Day 1	Nutrients	1	1.02	0.09	0.445	0.04	0.855
	Glyphosate	2	0.59	0.10	0.925	0.94	0.412
	Imidacloprid	2	0.76	0.13	0.775	1.86	0.182
	Nutrients and glyphosate	2	0.63	0.11	0.894		
	Nutrients and imidacloprid	2	0.49	0.09	0.979		
	Glyphosate and imidacloprid	2	1.15	0.20	0.362		
	Nutrients, glyphosate and imidacloprid	2	0.59	0.10	0.937		
Day 7	Nutrients	1	2.65	0.07	0.131	0.02	0.873
	Glyphosate	**2**	**10.40**	**0.58**	***0.005****	0.29	0.754
	Imidacloprid	2	1.13	0.06	0.429	3.21	0.065
	Nutrients and glyphosate	2	1.24	0.07	0.378		
	Nutrients and imidacloprid	2	1.06	0.06	0.461		
	Glyphosate and imidacloprid	2	0.96	0.05	0.472		
	Nutrients, glyphosate and imidacloprid	2	0.80	0.04	0.572		
Day 15	Nutrients	1	4.75	0.12	0.027	0.11	0.701
	Glyphosate	**2**	**10.06**	**0.53**	***0.001****	0.67	0.523
	Imidacloprid	2	0.91	0.05	0.537	0.93	0.930

	Nutrients and glyphosate	2	1.95	0.10	0.169		
	Nutrients and imidacloprid	2	0.92	0.05	0.532		
	Glyphosate and imidacloprid	2	1.12	0.06	0.417		
	Nutrients, glyphosate and imidacloprid	2	0.83	0.04	0.596		
Day 30	Nutrients	1	2.44	0.12	0.055	0.04	0.848
	Glyphosate	2	2.92	0.29	0.007	2.06	0.169
	Imidacloprid	2	1.11	0.11	0.365	0.84	0.425
	Nutrients and glyphosate	2	1.11	0.11	0.383		
	Nutrients and imidacloprid	2	1.19	0.12	0.300		
	Glyphosate and imidacloprid	2	0.89	0.09	0.590		
	Nutrients, glyphosate and imidacloprid	2	0.71	0.07	0.792		
Day 43	Nutrients	1	3.29	0.11	0.036	0.09	0.799
	Glyphosate	**2**	**6.42**	**0.44**	***0.001****	6.05	0.015*
	Imidacloprid	2	1.11	0.08	0.402	0.13	0.877
	Nutrients and glyphosate	2	1.55	0.10	0.217		
	Nutrients and imidacloprid	2	1.32	0.09	0.311		
	Glyphosate and imidacloprid	2	0.89	0.06	0.543		
	Nutrients, glyphosate and imidacloprid	2	0.81	0.05	0.637		

Table II.S7 PERMANOVA for different explanatory variables (and their interaction) in models with the Jensen-Shannon divergence among communities as the response. The same model was tested at five different time points and an asterisk indicates p-values that are significant after a Bonferroni correction for 7 hypothesis tests (i.e. only p-values <0.007 are considered significant, shown in bold). A PERMDISP was performed to confirm homogeneity of groups dispersions and significant *p*-values (<0.05) point out to predictors whose significance in the PERMANOVA output should be carefully analysed as they may be a result of within-group variation rather than among-group variation. df=degrees of freedom

			PERMANOVA (999 permutations)			PERMDISP (999 permutations)	
	Predictor	df	F	R^2	*p*-value	F	*p*-value
Day 1	Nutrients	1	1.37	0.11	0.166	0.26	0.260
	Glyphosate	2	0.55	0.09	0.972	0.57	0.592
	Imidacloprid	2	0.90	0.15	0.614	0.78	0.500
	Nutrients and glyphosate	2	0.71	0.11	0.874		
	Nutrients and imidacloprid	2	0.52	0.08	0.990		
	Glyphosate and imidacloprid	2	1.11	0.18	0.330		
	Nutrients, glyphosate and imidacloprid	2	0.69	0.11	0.883		
Day 7	Nutrients	1	2.03	0.11	0.080	0.16	0.701
	Glyphosate	**2**	**3.28**	**0.34**	***0.005****	1.52	0.269
	Imidacloprid	2	1.21	0.13	0.278	3.54	0.057
	Nutrients and glyphosate	2	0.93	0.10	0.512		
	Nutrients and imidacloprid	2	0.72	0.07	0.776		
	Glyphosate and imidacloprid	2	0.87	0.09	0.579		
	Nutrients, glyphosate and imidacloprid	2	0.63	0.07	0.858		
Day 15	Nutrients	1	3.29	0.12	0.027	0.00	0.979
	Glyphosate	**2**	**6.37**	**0.45**	***0.001****	2.13	0.165
	Imidacloprid	2	0.91	0.06	0.526	1.27	0.305
	Nutrients and glyphosate	2	1.73	0.12	0.136		

	Nutrients and imidacloprid	2	0.80	0.06	0.621		
	Glyphosate and imidacloprid	2	1.14	0.08	0.359		
	Nutrients, glyphosate and imidacloprid	2	0.68	0.05	0.759		
Day 30	Nutrients	1	2.99	0.12	0.013	0.12	0.737
	Glyphosate	**2**	**4.21**	**0.34**	***0.001****	8.02	0.004*
	Imidacloprid	2	1.48	0.12	0.125	1.02	0.378
	Nutrients and glyphosate	2	1.04	0.08	0.409		
	Nutrients and imidacloprid	2	1.07	0.09	0.385		
	Glyphosate and imidacloprid	2	1.18	0.09	0.297		
	Nutrients, glyphosate and imidacloprid	2	0.98	0.08	0.479		
Day 43	Nutrients	1	2.30	0.11	0.051	0.38	0.813
	Glyphosate	**2**	**3.85**	**0.37**	***0.001****	10.41	0.004*
	Imidacloprid	2	0.91	0.09	0.549	1.14	0.345
	Nutrients and glyphosate	2	0.96	0.09	0.535		
	Nutrients and imidacloprid	2	1.01	0.10	0.446		
	Glyphosate and imidacloprid	2	0.76	0.07	0.739		
	Nutrients, glyphosate and imidacloprid	2	0.78	0.07	0.699		

Table II.S8 PERMANOVA for glyphosate as the explanatory variable in models with weighted UniFrac distance or Jensen-Shannon divergence among communities as the response variable after data transformation by ASV relative abundance calculation (unrarefied) or by rarefying samples to 10,000 reads. The same model was tested at five different time points and an asterisk indicates p-values that are significant after a conservative Bonferroni correction for 7 hypothesis tests (i.e. only p-value<0.007 are considered significant). A PERMDISP was performed to confirm homogeneity of groups dispersions and significant p-values (<0.05) point out to predictors whose significance in the PERMANOVA output should be carefully analysed as they may be a result of within-group variation rather than among-group variation.

			Glyphosate effect (PERMANOVA 999 permutations)			Dispersion within glyphosate groups (PERMDISP 999 permutations)	
		Data transformation	F	R^2	*p*-value	F	*p*-value
Weighted UniFrac distance	Day 1	Relative abundance	1.23	0.10	0.357	0.95	0.420
		Reads rarefied to 10,000	1.24	0.10	0.358	0.71	0.522
	Day 7	Relative abundance	3.68	0.41	0.040	0.52	0.585
		Reads rarefied to 10,000	3.31	0.42	0.060	0.74	0.495
	Day 15	Relative abundance	**6.06**	**0.33**	***0.005****	0.01	0.991
		Reads rarefied to 10,000	**6.00**	**0.33**	***0.002****	0.00	0.996
	Day 30	Relative abundance	2.10	0.19	0.060	0.70	0.505
		Reads rarefied to 10,000	1.48	0.21	0.239	0.64	0.518
	Day 43	Relative abundance	2.83	0.36	0.031	**7.14**	***0.012****
		Reads rarefied to 10,000	4.11	0.44	0.045	**4.94**	***0.035****
Jensen-Shannon divergence	Day 1	Relative abundance	0.65	0.08	0.831	0.60	0.522
		Reads rarefied to 10,000	0.65	0.08	0.833	0.66	0.530
	Day 7	Relative abundance	3.36	0.38	0.039	0.14	0.881
		Reads rarefied to 10,000	3.10	0.40	0.059	0.43	0.682
	Day 15	Relative abundance	**6.66**	**0.37**	***0.002****	0.13	0.863
		Reads rarefied to 10,000	**6.59**	**0.37**	***0.002****	0.14	0.872
	Day 30	Relative abundance	3.92	0.29	0.007	1.41	0.273
		Reads rarefied to 10,000	2.87	0.30	0.056	1.23	0.298

	Day 43	Relative abundance	**3.54**	**0.36**	***0.005****	**4.51**	***0.035****
		Reads rarefied to 10,000	3.23	0.41	0.036	1.90	0.189

Table II.S9 Percent of variance explained by the two first PRC axes, and by time or treatment when nutrient treatments are grouped as replicates (see Figure II.6 and Figure II.S5). F statistic and p-value of permutation test for first constrained eigenvalue is also shown, and an asterisk denote significant p-values.

Taxonomic level	**PRC1 (%)**	**PRC2 (%)**	**Time (%)**	**Time and treatment (%)**	**Permutation test (999 permutations)** F	*p*-value
Phylum	47.7	15.3	15.1	46.5	31.22	0.001*
Class	46.4	12.6	13.8	49.8	34.28	0.001*
Order	38.6	14.5	15.3	47.2	26.19	0.001*
Family	33.5	11.9	16.3	45.2	21.30	0.001*
Genus	34.9	10.6	14.4	44.7	20.6	0.001*
ASV	22.1	8.0	11.1	41.9	10.61	0.001*

Table II.S10 All bacterioplankton taxa weights for the PRC model at the phylum level, ranked from largest (positive effects of glyphosate treatment) to smallest (negative effects of glyphosate treatment).

Phylum	**Weight in PRC1**
Proteobacteria	0.37
Chlamydiae	0.07
Acidobacteria	0.04
GN02	0.02
Firmicutes	0.01
BRC1	0.01
SR1	0.004
Fibrobacteres	2.33E-27
Lentisphaerae	2.17E-36
Nitrospirae	4.23E-39
WS3	0.00
OP3	-3.23E-44
Fusobacteria	-1.14E-30
Tenericutes	-0.002
TM6	-0.03
[Thermi]	-0.03
Gemmatimonadetes	-0.04
TM7	-0.05
Spirochaetes	-0.05
Chlorobi	-0.06
Armatimonadetes	-0.08
Chloroflexi	-0.08
OD1	-0.09
Verrucomicrobia	-0.12
Planctomycetes	-0.14
Cyanobacteria	-0.24
Actinobacteria	-0.26
Bacteroidetes	-0.43

Table II.S11 Relative abundance of the main affected phyla by treatment. Percentage was calculated after normalization with DESeq2 or by rarefying samples to 10,000 reads each and the respective standard error is indicated in parenthesis.

Normalization	Treatment	Proteobacteria	Bacteroidetes	Actinobacteria	Cyanobacteria
DESeq2	Glyphosate 15 mg/L and imidacloprid 60 ug/L	62.1% (3.2)	11.7% (2.1)	2.6% (0.8)	3.3% (0.8)
	Glyphosate 15 mg/L	68.2% (3.7)	10.3% (1.8)	3.9% (1.5)	1.6% (0.6)
	Imidacloprid 60 ug/L	48.4% (1.4)	20.7% (0.9)	7.0% (1.0)	3.6% (0.6)
	Glyphosate 0.3 mg/L and imidacloprid 1 ug/L	46.4% (0.9)	20.1% (0.9)	5.9% (0.8)	4.0% (0.6)
	Glyphosate 0.3 mg/L	46.9% (1.4)	20.6% (1.4)	6.5% (1.2)	4.6% (1.7)
	Imidacloprid 1 ug/L	48.9% (1.3)	20.1% (1.4)	6.4% (1.1)	3.4% (0.5)
	Control	49.6% (1.2)	22.4% (1.1)	5.2% (0.8)	3.4% (0.6)
Rarefied to 10k reads	Glyphosate 15 mg/L and imidacloprid 60 ug/L	69.5% (5.7)	12.6% (3.9)	3.2% (1.6)	1.1% (0.4)
	Glyphosate 15 mg/L	66.7% (7.7)	8.5% (2.5)	9.5% (5.5)	0.7% (0.5)
	Imidacloprid 60 ug/L	52.8% (4.0)	17.8% (3.9)	19.9% (4.4)	1.2% (0.4)
	Glyphosate 0.3 mg/L and imidacloprid 1 ug/L	38.6% (2.8)	20.3% (2.6)	15.8% (4.2)	2.9% (1.6)
	Glyphosate 0.3 mg/L	39.8% (3.4)	16.5% (4.4)	14.4% (5.7)	8.1% (4.9)
	Imidacloprid 1 ug/L	44.3% (1.7)	24.9% (2.7)	15.3% (4.2)	1.9% (0.5)
	Control	44.4% (4.1)	20.3% (2.5)	12.7% (3.9)	2.2% (0.8)

Table II.S12 ASVs with the highest PRC taxa weights, and their respective weight in the first RDA axis, the ratio between this value and the maximum taxa weight of the PRC model, and their taxonomy assignment from TaxAss using FreshTrain and GreenGenes databases.

ASV	Weight in RDA1	(Taxa weight) / (max. weight)	Phylum	Class	Order	Family/Lineage	Genus/Clade	Species/Tribe
sp2262	0.14	1.00	Proteobacteria	Alphaproteobacteria	unclassified	unclassified	unclassified	unclassified
sp188	0.13	0.92	Proteobacteria	Betaproteobacteria	Burkholderiales	betI	betI-A	Lhab-A4
sp1895	0.12	0.88	Proteobacteria	Alphaproteobacteria	Rhodospirillales	Rhodospirillaceae	*Azospirillum*	*A. massiliensis*
sp2118	0.12	0.84	Proteobacteria	Alphaproteobacteria	Rhodobacterales	alfVI	unclassified	unclassified
sp2155	0.11	0.81	Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiaceae	*Agrobacterium*	unclassified
sp307	0.11	0.81	Proteobacteria	Gammaproteobacteria	Xanthomonadales	Sinobacteraceae	*Nevskia*	*N. ramosa*
sp284	0.10	0.75	Proteobacteria	Betaproteobacteria	Burkholderiales	betI	betI-A	unclassified
sp130	0.10	0.74	Proteobacteria	Betaproteobacteria	Methylophilales	betIV	betIV-A	unclassified
sp283	-0.11	-0.80	Proteobacteria	Betaproteobacteria	Burkholderiales	betI	betI-A	unclassified
sp2111	-0.11	-0.82	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	*Rhodobacter*	unclassified

Table II.S13 Summarized results of the generalized additive mixed models (GAMMs) for abundance of the three genera most positively impacted by the experimental treatments. Gaussian residual distributions were used in all models. For each response variable we report the sample size (n), adjusted R2, the predictors and factors used in the model, the parameter estimate and respective standard error (SE) of parametric effects or the effective degrees of freedom (EDF) of smooth terms, the corresponding test statistics (t value for parametric and F for smooth terms) and the p-value. Smooths terms are described as mgcv syntax: 's()' functions are thin plate regression splines and 't()' tensor product interactions, pond represents the random variable of the mixed model and 'bs='fs'' specifies the underline base function as a random smooth. Following a Bonferroni multiple testing correction for 16 tests, we only considered significant variables with p-value <0.003, shown in bold.

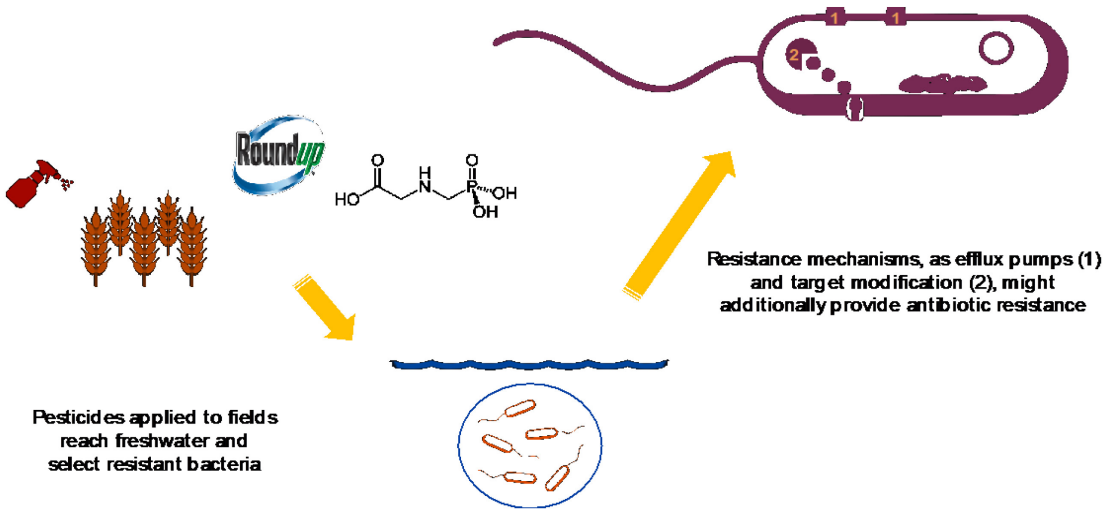
Response variable	n	Adjusted R^2	Predictors	Factors† of parametric and smooth terms	Estimate (SE) or EDF	Statistic	p-value
log_{10}(*Agrobacterium*)	89	0.80	**treatment**	**high_both**	**1.26 (0.17)**	**7.50**	**<0.001***
				high_glypho	**1.05 (0.17)**	**6.25**	**<0.001***
				high_imid	0.06 (0.17)	0.38	0.707
				low_both	0.26 (0.17)	1.52	0.134
				low_glypho	0.44 (0.19)	2.37	0.021
				low_imid	0.15 (0.17)	0.87	0.389
			nutrient	high	0.10 (0.10)	0.99	0.328
			ti(day)		4.8	2.92	0.021
			ti(day, by=treatment)	low_glypho	1.4	5.43	0.007
				low_imid	1.0	1.05	0.309
				low_both	1.0	0.57	0.453
				high_glypho	**3.9**	**19.49**	**<0.001***
				high_imid	1.0	0.07	0.793
				high_both	**3.9**	**20.66**	**<0.001***
			ti(day, by=nutrient)	high	1.0	0.01	0.941
			s(day, Pond, bs='fs')		3.0	0.14	0.24
log_{10}(*Flavobacterium*)	89	0.48	**treatment**	**high_both**	**1.5 (0.25)**	**5.96**	**<0.001***
				high_glypho	**1.0 (0.25)**	**3.98**	**<0.001***
				high_imid	0.24 (0.25)	0.95	0.348
				low_both	0.46 (0.25)	1.82	0.074

				low_glypho	0.51 (0.28)	1.80	0.076
				low_imid	0.16 (0.15)	0.64	0.522
			nutrient	high	-0.06 (0.15)	-0.40	0.687
			ti(day)		4.7	2.78	0.042
			ti(day, by=treatment)	low_glypho	1.8	1.17	0.288
				low_imid	1.0	2.24	0.139
				low_both	1.9	1.01	0.325
				high_glypho	2.1	3.63	0.031
				high_imid	1.0	0.27	0.606
				high_both	**1.0**	**17.35**	**<0.001***
			ti(day, by=nutrient)	high	2.0	0.64	0.442
			s(day, Pond, bs='fs')		0.2	0.01	0.421
log_{10}(*Azospirillum*)	89	0.78	**treatment**	**high_both**	**1.34 (0.17)**	**7.82**	**<0.001***
				high_glypho	**1.02 (0.17)**	**5.99**	**<0.001***
				high_imid	0.09 (0.17)	0.50	0.62
				low_both	0.31 (0.17)	1.83	0.07
				low_glypho	0.18 (0.18)	0.98	0.33
				low_imid	0.18 (0.17)	1.06	0.29
			nutrient	high	-0.24 (0.10)	-2.47	0.02
			ti(day)		2.2	2.11	0.137
			ti(day, by=treatment)	low_glypho	1.0	0.01	0.920
				low_imid	1.0	0.00	0.988
				low_both	1.6	0.76	0.480
				high_glypho	**2.6**	**5.41**	**0.002***
				high_imid	1.0	0.09	0.770
				high_both	**3.3**	**6.27**	**0.001***
			ti(day, by=nutrient)	high	2.2	1.40	0.192
			s(day, Pond, bs='fs')		7.8	0.60	0.015

CHAPTER III : A glyphosate-based herbicide cross-selects for antibiotic resistance genes in bacterioplankton communities

Naíla Barbosa da Costa[1,2], Marie-Pier Hébert[2,3] , Vincent Fugère[2,4,5], Yves Terrat[1], Gregor F. Fussmann[2,3,4], Andrew Gonzalez[3,4], B. Jesse Shapiro[1,2,4,6,7]

[1] Département des sciences biologiques, Université de Montréal, Montreal, Canada; [2] Groupe de Recherche Interuniversitaire en Limnologie et environnement aquatique (GRIL), Montreal, Canada; [3] Department of Biology, McGill University, Montreal, Canada; [4] Québec Centre for Biodiversity Science (QCBS), Montreal, Canada; [5] Département des sciences de l'environnement, Université du Québec à Trois-Rivières, Trois-Rivières, Canada; [6] Department of Microbiology and Immunology, McGill University, Montreal, Canada; [7] McGill Genome Centre, McGill University, Montreal, Canada



Published in *mSystems*, 7(2): e01482-21. 10.1128/msystems.01482-21

Minor edits to the published text have been made here, following suggestions by this thesis' reviewers

Abstract

Agrochemicals often contaminate freshwater bodies, affecting microbial communities that underlie aquatic food webs. For example, the herbicide glyphosate has the potential to indirectly select for antibiotic resistant bacteria. Such cross-selection could occur, for example, if the same genes (encoding efflux pumps, for example) confer resistance to both glyphosate and antibiotics. To test for cross-resistance in natural aquatic bacterial communities, we added a glyphosate-based herbicide (GBH) to 1,000-L mesocosms filled with water from a pristine lake. Over 57 days, we tracked changes in bacterial communities with shotgun metagenomic sequencing, and annotated metagenome-assembled genomes (MAGs) for the presence of known antibiotic resistance genes (ARGs), plasmids, and resistance mutations in the enzyme targeted by glyphosate (enolpyruvyl-shikimate-3-phosphate synthase; EPSPS). We found that high doses of GBH significantly increased ARG frequency and selected for multidrug efflux pumps in particular. The relative abundance of MAGs after a high dose of GBH was predictable based on the number of ARGs encoded in their genomes (17% of variation explained) and, to a lesser extent, by resistance mutations in EPSPS. Together, these results indicate that GBHs have the potential to cross-select for antibiotic resistance in natural freshwater bacteria.

Importance

Glyphosate-based herbicides (GBHs) such as “Roundup®” formulations may have the unintended consequence of selecting for antibiotic resistance genes (ARGs), as demonstrated in previous experiments. However, the effects of GBHs on ARGs remains unknown in natural aquatic communities, which are often contaminated with pesticides from agricultural runoff. Moreover, the resistance provided by ARGs compared to canonical (i.e. previously known) mutations in the glyphosate target enzyme, EPSPS, remains unclear. Here we performed a freshwater mesocosm experiment showing that a GBH strongly selects for ARGs, particularly multidrug efflux pumps. These selective effects were evident after just a few days, and the ability of bacteria to survive and thrive after GBH stress was predictable by the number of ARGs in their genomes, and to a

lesser extent by mutations in EPSPS. Intensive GBH application may therefore have the unintended consequence of selecting for ARGs in natural freshwater communities.

Introduction

Glyphosate-based herbicides (GBHs) are by far the most extensively used weed-killers worldwide, especially since the introduction of transgenic glyphosate-resistant crops in the 1990s (Benbrook, 2016; Duke & Powles, 2008). Glyphosate residues can spread widely and accumulate in soil, water, and plant products, raising concerns over human and environmental health (van Bruggen et al., 2018). A recent systematic review and risk analysis concluded that glyphosate poses a moderate to high risk to freshwater biodiversity in 20 of the countries investigated (Brovini et al., 2021). Some of the highest aquatic concentrations of glyphosate were found in countries with the largest production of genetically engineered glyphosate-tolerant crops globally, including the United States, Brazil, and Argentina (Benbrook, 2016; Brovini et al., 2021).

Although designed to control weed growth, glyphosate may also affect microorganisms that use the herbicide's molecular target, the enzyme enolpyruvyl-shikimate-3-phosphate synthase (EPSPS), to synthesize aromatic amino acids (Pollegioni et al., 2011). The EPSPS is classified into four classes according to mutations in the enzyme active site that confer differential sensitivities to glyphosate (Leino et al., 2020). In bacteria, EPSPS classes I and II, which are respectively sensitive and tolerant to glyphosate, are the most frequently found, while classes III and IV are rarer and both confer glyphosate resistance (Leino et al., 2020). The EPSPS class II sequence isolated from a strain of *Agrobacterium tumefaciens* is used as the transgene conferring tolerance in most commercially available glyphosate-resistant crops (Pline-Srnic, 2006; Singh & Prasad, 2016).

Experiments conducted in diverse environments, such as soil and freshwater (Barbosa da Costa et al., 2021; Lu et al., 2020; Y. Wang et al., 2020) and the bee gut microbiome (Motta, Raymann, & Moran, 2018), have shown that bacterial taxa from natural ecosystems vary in their sensitivity to glyphosate. Some of this variation is explained by the distribution of different EPSPS classes. However, while strains with the EPSPS class I are known to be sensitive, they have also been observed to tolerate

glyphosate through unknown mechanisms (Motta et al., 2018), indicating that additional EPSPS-independent glyphosate resistance mechanisms likely exist in nature.

Studies with bacterial cultures have shown increased resistance to antibiotics after exposure to high concentrations of glyphosate and other herbicides (Kurenbach et al., 2017, 2018, 2015; Ramakrishnan, Venkateswarlu, Sethunathan, & Megharaj, 2019; Xing, Wu, & Men, 2020). In the presence of glyphosate, the expression of membrane transporters may confer resistance to glyphosate and antibiotics simultaneously (Staub, Brand, Tran, Kong, & Rogers, 2012). Specifically, multidrug efflux pumps have been experimentally shown to confer resistance to both glyphosate and antibiotics, presumably by exporting a variety of small molecules (Kurenbach et al., 2017, 2015). This is an example of cross-resistance, a mechanism of indirect selection through which one resistance gene or biochemical system confers resistance to other antimicrobial agents (Baker-Austin, Wright, Stepanauskas, & McArthur, 2006; Murray, Zhang, Snape, & Gaze, 2019).

Direct selection of antibiotic resistance occurs when bacteria are exposed to an antibiotic agent and mutations conferring resistance to this agent are selected (Gullberg et al., 2011). In contrast, indirect selection for antibiotic resistance occurs in the absence of the antibiotic, either via cross- or co-resistance (Baker-Austin et al., 2006; Murray et al., 2019). Cross-resistance occurs when the same gene confers resistance to multiple antibiotic agents, while co-resistance occurs when a resistance gene is genetically linked to another gene that is not necessarily an antibiotic resistance gene (ARG), but that is under positive selection.

Most studies of cross-resistance induced by herbicides have focused on bacterial isolates in laboratory experiments (Kurenbach et al., 2017, 2018, 2015; Xing et al., 2020; H. Zhang, Liu, Wang, & Zhai, 2021). A recent study showed that herbicide selection increases the prevalence of ARGs in soil bacterial communities, using observational and experimental field data (Liao et al., 2021). However, we still lack evidence in aquatic communities, which are of particular interest because herbicides often reach water bodies through leaching, runoff, and spray drift from agricultural fields (Brook & Beaton, 2015; Brovini et al., 2021). Moreover, the extent of direct selection on EPSPS mutations compared to indirect selection on ARGs is unclear. In a previous study, we used 16S

ribosomal gene amplicon sequencing to assess how the composition of freshwater bacterioplankton communities respond to a GBH applied alone or in combination with a widely-used neonicotinoid insecticide (Barbosa da Costa et al., 2021). As part of the same experiment, we also showed that phytoplankton and zooplankton communities responded strongly to high doses of GBH (Fugère et al., 2020; M.-P. Hébert et al., 2021). Here, we expand on our previous work and investigate the effects of the GBH on ARG frequencies in aquatic bacterial communities, using the same outdoor array of experimental ponds (Figure III.1A).

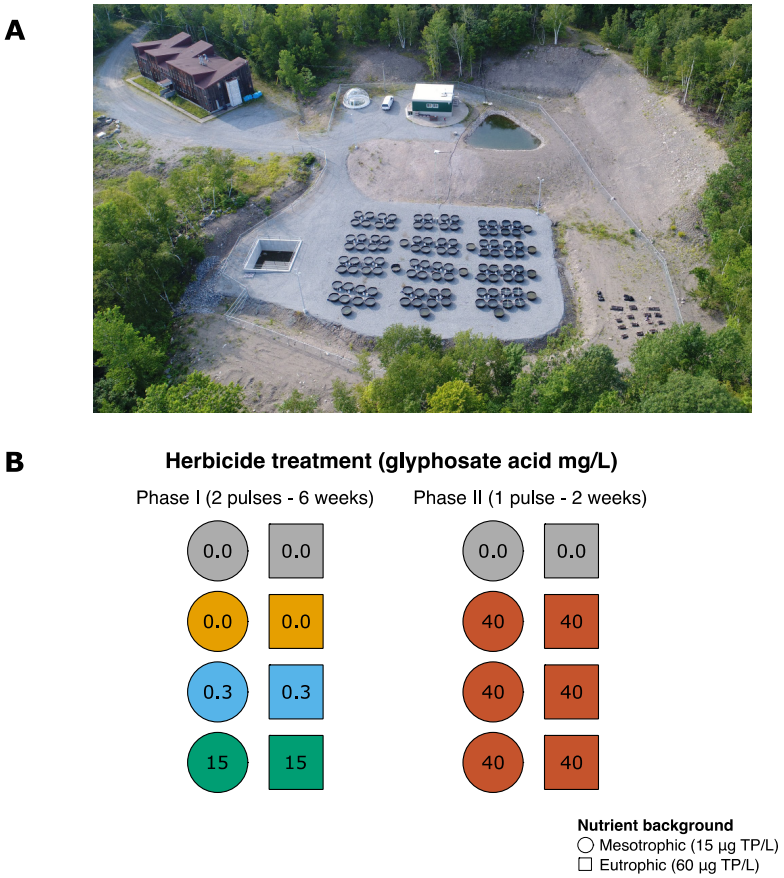


Figure III.1 Experimental area and design. (A) Aerial photograph of the Large Experimental Array of Ponds (LEAP) at Gault Nature Reserve, in Mont Saint-Hilaire (Canada). The laboratory facility and inflow reservoir, where water from our source lake was redirected to before filling the mesocosms, can be seen at the top of the photograph. Our source lake, Lake Hertel, is located upstream (not shown in the photograph). (B) Schematic representation of the subset of mesocosms selected for metagenomic sequencing in this study. A total of eight ponds were sampled 11 times over the course of the 8-week experiment, which was divided in two phases: Phase I (6 weeks) and Phase II (2 weeks). Phase I included two pulse applications (doses) of GBH, with three target glyphosate concentrations (0, 0.5, and 15 mg/L). In Phase II, all ponds except for two

controls, shown in grey, received a higher dose of glyphosate (40 mg/L). Phase I included four control ponds (grey and yellow) while Phase II only included two controls (grey). Note that yellow ponds only received GBH in Phase II. Nutrients were also added to ponds to reproduce mesotrophic or eutrophic conditions, represented respectively by circles and squares (target phosphorus concentrations are indicated). TP: total phosphorus.

To test the extent to which GBH (in the form of Roundup®) cross-selects for ARGs in complex aquatic communities over time, we exposed freshwater mesocosms to two glyphosate concentrations for six weeks (0.3 and 15 mg/L; Phase I) and to a higher dose for the next two weeks (40 mg/L; Phase II) (Figure III.1B). We sequenced metagenomes from each mesocosm and reconstructed Metagenome-Assembled Genomes (MAGs) of bacteria, which were annotated according to their taxonomy, presence of ARGs, plasmids, and resistance mutations in the EPSPS enzyme (Methods). We hypothesize that the frequency of ARGs would increase after exposure to a high concentration of glyphosate, and that efflux pumps are among the main resistance mechanisms promoted by GBH. We also expect that MAGs encoding many ARGs and/or the resistant classes of the EPSPS gene will be the most likely to survive and proliferate after GBH exposure. Consistent with these expectations, we find that high doses of GBH (15 and 40 mg/L glyphosate) cross-select for ARGs, particularly multidrug efflux pumps. These results demonstrate that severe contamination of aquatic systems with GBH could indirectly select for antibiotic resistance.

Methods

Experimental design

An eight-week mesocosm experiment was conducted at the Large Experimental Array of Ponds (LEAP) facility (Figure III.1A) located at McGill University's Gault Nature Reserve (QC, Canada) from August 17th (day 1) to October 12th (day 57) 2016, as previously described (Barbosa da Costa et al., 2021; Fugère et al., 2020; M.-P. Hébert et al., 2021). Pond mesocosms were filled with 1,000 L of water and planktonic communities from Lake Hertel (45°32' N, 73°09' W). Lake water was passed through a coarse sieve to prevent fish introduction, while retaining lake bacterioplankton, zooplankton and phytoplankton, whose responses to experimental treatments have been described in

previous studies (Barbosa da Costa et al., 2021; Fugère et al., 2020; M.-P. Hébert et al., 2021).

Figure III.1B illustrates the experimental design of a subset of eight treatments selected for the metagenomic sequencing analyses reported here (see Fugère et al. (2020) for a full description of all treatments at the LEAP facility in 2016). The eight ponds were sampled at 11 timepoints throughout phases I and II of the experiment. In Phase I (days 1-44), all ponds received nutrient inputs biweekly, simulating mesotrophic or eutrophic lake conditions with additions of a concentrated nutrient solution. Four ponds were treated with a GBH to reach target concentrations of 0.3 or 15 mg/L of the active ingredient (glyphosate; acid equivalent), while the other four were kept as control ponds. The GBH was applied in two pulses in Phase I, at days 6 and 33. In Phase II (days 45-57), two control ponds (hereafter referred to as Control Phase I) and the four treatment ponds received one pulse of the GBH at a higher dose (40 mg/L glyphosate) on day 44, while other two other control ponds (hereafter referred to as Control Phase II) received no pulse.

Target doses of the active ingredient were calculated based on the glyphosate acid content in Roundup® Super Concentrate Grass and Weed Control (Bayer ©; reg. #22759), the formulation used for the experiment. We used a commercial formulation to mimic environmental contamination, and because the costs of using pure glyphosate salt would be prohibitive in a large-scale field experiment. Treatments are referred to by their glyphosate acid concentration to allow comparison with other formulations. Nutrients were added in the form of nitrate (KNO_3) and phosphate (KH_2PO_4 and K_2PO_4), with target concentrations of 15 µg P/L and 231 µg N/L in the low-nutrient (mesotrophic) treatment ponds and 60 µg P/L and 924 µg N/L for in the high-nutrient (eutrophic) treatment ponds. The concentrated nutrient solution had an N:P molar ratio of 33 comparable to our source lake. As reported in previous studies (Barbosa da Costa et al., 2021; Fugère et al., 2020), target doses of glyphosate acid and nutrients were achieved reasonably well, although glyphosate accumulated over Phase I, reaching higher concentrations than intended after the second dose.

DNA extraction and metagenomic sequencing

The eight experimental ponds were sampled for bacterioplankton DNA at 8 timepoints during Phase I (days 1, 7, 15, 30, 35, 38, 41 and 43) and 3 timepoints during Phase II (days 45, 49 and 57). Water samples were collected with 35 cm long integrated samplers (2.5 cm diameter PVC tubing) at multiple locations in the same pond and stored in 1 L dark Nalgene bottles, at 4 °C until being filtered within 4 hours. We filtered 250 mL of each sample on site, through 0.22 µm pore size Millipore hydrophilic polyethersulfone membranes of 47 mm diameter (Sigma-Aldrich, St. Louis, USA). Filters were stored at -80 ºC until DNA extraction.

We extracted DNA from a total of 88 filter samples using the PowerWater DNA Isolation kit (MoBio Technologies Inc.) following the manufacturer's guidelines. Shotgun metagenomic sequencing was performed using the Illumina HiSeq 4000 technology with 100 bp paired-end reads. Libraries were prepared with 50 ng of DNA using the NEBNext Ultra II DNA Library Prep kit for Illumina (New England Biolabs®) as per the manufacturer's recommendations, and had an average fragment size of 390 bp.

Metagenomic read trimming, functional annotation and ARGs inference from metagenomic reads

We removed Illumina adapters and quality filtered metagenomic reads using Trimmomatic (Bolger, Lohse, & Usadel, 2014) in the paired-end mode. We used FragGeneScan (Rho, Tang, & Ye, 2010) for gene prediction from trimmed metagenomic reads and annotated predicted genes with SEED subsystems (Overbeek et al., 2014). To identify known ARGs in the metagenomic reads, we used the Resistance Gene Identifier (RGI) 'bwt' function that maps FASTQ files of reads passing quality control to CARD (Alcock et al., 2020) using Bowtie2 (version 2.4) as an aligner (Langmead & Salzberg, 2012). Only alignments with mapping quality (MAPQ) higher than 10 and gene coverage of 50% were retained. To calculate the proportion of metagenomic reads mapped to CARD that have been assembled and binned to genomes, we extracted reads that aligned to CARD using Samtools (H. Li et al., 2009) and mapped them to MAGs using Bowtie2 (Langmead & Salzberg, 2012). Table III.S2 shows the total number of reads by sample after trimming and a summary of the RGI output by sample for hits with minimum gene coverage of 50% and average MAPQ>10.

Metagenomic de novo co-assembly, binning, dereplication and curation of MAGs

We organized the dataset into eight sets of metagenomes, each of them containing samples of the same mesocosm pond (Figure III.1B) from multiple timepoints. We co-assembled reads from each of the 8 timeseries using MEGAHIT v1.1.1 (D. Li, Liu, Luo, Sadakane, & Lam, 2015), with a minimum contig length of 1 kbp. We used anvi'o v5.1 (Eren et al., 2015) to profile contigs, to identify genes using Prodigal v2.6.3 (Doug Hyatt, Gwo-Liang Chen, Philip F LoCascio, Miriam L Land, Frank W Larimer, 2010) and HMMER v3.2.1 (Eddy, 2011), to infer the taxonomy of genes with Centrifuge v1.0.4 (Kim, Song, Breitwieser, & Salzberg, 2016), to map metagenomic reads to contigs using Bowtie2 v2.4.2 (Langmead & Salzberg, 2012), and then to estimate depth of read coverage across contigs. Finally, we used anvi'o to cluster contigs according to their sequence composition and coverage across samples with the automatic binning algorithm CONCOCT (Alneberg et al., 2014) and we manually refined the bins (*n*=830) using the anvi'o interactive interface, as suggested by developers (Eren et al., 2015), by removing splits that diverged in the differential coverage and/or tetra-nucleotide frequency of most splits in the same bin.

We dereplicated bins as described in Delmont et al. (2018). In summary, we calculated the Pearson correlation coefficient between the relative abundance (i.e. the mean coverage calculated by the function 'anvi-summarize' within anvi'o) for each pair of bins in the metagenomic samples, using the 'cor' function in R (R Core Team, 2020), and the average nucleotide identity (ANI) of bins affiliated to the same phylum, using NUCmer (Delcher, Phillippy, Carlton, & Salzberg, 2002). Taxonomy assignment of redundant bins was done using CheckM (Parks, Imelfort, Skennerton, Hugenholtz, & Tyson, 2015). Bins with a Person correlation coefficient above 0.9 and ANI of 98% or more were considered redundant. In a total of 830 bins obtained before performing the dereplication, we found 607 non-redundant bins, of which 426 were classified as MAGs, as they had at least 70% completeness and no more than 10% redundancy (see Table III.S2). We then created a non-redundant genomic database of these 426 MAGs to which we mapped metagenomic reads to calculate the relative abundance of each MAGs across the different samples. Here we define a MAG's relative abundance as the number of metagenomic reads recruited to a MAG divided by the total metagenomic reads in a given sample.

Identification of ARGs, EPSPS and plasmids in MAGs

We annotated ARGs within MAG contigs with the RGI 'main' function, that compares predicted protein sequences from contigs to the CARD protein reference sequence data. Within RGI, we used the BLAST (Altschul, Gish, Miller, Myers, & Lipman, 1990) alignment option and the strict algorithm (excluding nudge of loose hits to strict hits) for low quality contigs (<20,000 bp). The RGI low sequence quality option uses Prodigal anonymous mode (Doug Hyatt, Gwo-Liang Chen, Philip F LoCascio, Miriam L Land, Frank W Larimer, 2010) for the prediction of open reading frames, supporting calls of partial ARGs from short or low quality contigs.

To identify EPSPS sequences from MAG contigs we first used anvi'o to predict amino acid sequences of the non-redundant MAGs with the flag 'report-aa-seqs-for-gene-calls' of the function 'anvi-summarize'. Gene calls of all the MAGs were concatenated conserving the original split names, and transformed into a fasta file. We then blasted the predicted amino acid sequences against a custom database with sequences of the EPSPS enzyme, using BLASTp (Altschul et al., 1990) and a minimum e-value of 1e-5. After selecting the gene call with the best match (i.e. lowest e-value) to an EPSPS sequence in each of the 426 MAGs, we used the *EPSPSClass* web server (Leino et al., 2020) to classify the retrieved sequences according to resistance to glyphosate. Sequences were classified as EPSPS class I, class II or class IV if they contained all the amino acid markers from the respective reference, i.e. if the percent identity was equal to 1; and classified as class III when they contained at least one complete motif out of 18 of the resistance-associated sequences, as explained in Leino et al. (2020). MAGs whose EPSPS sequences did not match these criteria of having at least one motif of class III or 100% percent identity with class I, II or IV, or those in which no predicted amino acid sequence matched a known EPSPS sequence were set as unclassified (roughly 27% of MAGs). EPSPS sequences matching class I were considered as putative sensitive and those with at least one motif of class III or matching class II as putative resistant. No sequences were found that matched to class IV.

To identify potential plasmid contigs assembled to MAGs we used the plasmid classifier PlasClass (Pellow, Mizrahi, & Shamir, 2020). We counted all contigs classified as plasmid with a minimum of 70% probability, as well as how many of these potential

plasmid contigs were annotated with ARGs through RGI. Table III.S2 summarizes MAG information, including the predicted EPSPS sequence found in the genome, the EPSPS classification, the number of estimated plasmid contigs and how many of them contained ARG sequences.

Statistical analyses

All statistical analyses were conducted in R v.4.0.2 (R Core Team, 2020). Time series of (log-transformed) ARG counts and ARG reads per million metagenomic reads were modelled using additive models (GAM) using the 'mgcv' R package (Wood, 2017). We used GAMs to account for nonlinear relationships among the response variable and the predictors. Some predictors (nutrient and herbicide treatment levels) were coded as ordered factors; Table III.1 lists all factors and predictors of the model. We built the models using the 'gam' function and assessed significance of effects with the 'summary.gam' function. We validated the models with the 'gam.check' function, inspecting the distribution of model residuals, comparing fitted and observed values, and checking if the basis dimension (*k*) of smooth terms were large enough.

We used Principal Response Curves (PRCs) to test for the effect of treatments on the composition of MAGs, ARGs and gene functional profiles over time. PRCs are a special case of partial redundancy analysis (pRDA) used in temporal experimental studies where treatments and the interaction between treatment and time are used as explanatory variables (Auber et al., 2017). Time is the covariable (or conditioning variable) whose effect is partialled out and the response variable is the matrix containing compositional data (taxa or gene family relative abundances). We built PRCs using relative abundances of predicted genes grouped according to the SEED subsystem levels 1 and 2. In a more focused analysis, we built a PRC for the matrix of ARGs found in each sample, i.e. metagenomic reads mapped to each ARG from the CARD reference classified according to their Antibiotic Resistance Ontology (ARO). The matrices were transformed using the Hellinger transformation (Legendre & Gallagher, 2001). The PRC diagram displays the treatment effect on the y-axis, expressed as deviations from the experimental controls at each time point. It also shows species scores on the right y-axis, which here can be interpreted as the contribution of each function or gene to the treatment response curves. We assessed the significance of the first PRC axis by permuting the

treatment label of ponds while keeping the temporal order, using the ‘permute’ package (Simpson, 2019) followed by a permutation test (999 permutations) using the ‘vegan’ package (J. Oksanen et al., 2019). For the PRC based on ARG composition, we tested if the distribution of PRC positive and negative scores was different among the resistance mechanisms of the identified ARGs using the ‘fisher.test’ function in the ‘stats’ package in R (R Core Team, 2020).

To test if MAG abundance in Phase II glyphosate treatments was correlated with their antibiotic resistance potential, we built a multiple linear regression with the ‘lm’ function of the R package ‘stats’ (R Core Team, 2020). The response variable was the average relative abundance of a MAG in glyphosate-treated ponds in Phase II, averaging abundance across all ponds in which a MAG was found in Phase II. The three predictors were: the MAG's antibiotic resistance potential (defined as the number of RGI strict hits found in the MAG), the average MAG relative abundance in the same ponds of Phase I, and their EPSPS sequence classification (resistant, sensitive or unclassified). To assess the relative contribution of the different predictors to MAG survival in Phase II, we performed a variance partitioning analysis with the ‘varpart’ function of the R package ‘vegan’ (J. Oksanen et al., 2019). Finally, to visualize the hierarchy among predictors we constructed a conditional inference regression tree. Response variable and predictors were the same as described above, except that instead of grouping all ARG hits, we transformed them into three variables, according to their function: antibiotic target alteration, antibiotic inactivation, or antibiotic efflux. The regression tree was fitted with the ‘ctree’ function in the R package ‘party’ (R Core Team, 2020). As a negative control, we repeated the same analyses for MAGs found in control ponds of Phase II.

As multiple predictors were tested, we used a conservative Bonferroni correction for the additive and linear models, whereby the *p*-value significance threshold of 0.05 was divided by the number of statistical tests.

Graphs and heatmaps for timeseries data visualization were built using the functions ‘geom_point’ and ‘geom_tile’, respectively, in the R package ‘ggplot2’ (Wickham, 2009).

Results

GBH treatment changes community composition and increases ARG frequency

To assess how GBH treatments affected bacterioplankton community composition over time, we built Principal Response Curves (PRCs; Methods). The response variables used for PRCs were either the estimated relative abundance of MAGs or the summed abundance of MAGs grouped at more inclusive taxonomic levels (phylum or class). In Phase I of the experiment, two pulses of a GBH were applied to reach concentrations of 0.3 mg/L or 15 mg/L glyphosate. The first GBH pulse at the highest concentration changed the MAG composition irreversibly for the duration of the experiment (Figure III.S1A). However, when these MAGs were grouped at more inclusive taxonomic levels, we observed a recovery of community composition 20-30 days after the first pulse (Figure III.S1 B-C). The second pulse in Phase I had a weaker effect that depended on the taxonomic resolution, with the most pronounced changes visible at the class level. In Phase II, when a single dose of 40 mg/L glyphosate was applied to all mesocosms except for the Phase II controls, a strong effect was observed in all ponds regardless of taxonomic resolution. These results are broadly consistent with our previous 16S rRNA gene amplicon sequencing from the same experiment, which showed community resilience to GBH pulses at higher taxonomic levels only (Barbosa da Costa et al., 2021).

To test the effects of GBH on ARG frequency over the experiment, we tracked variation in the number of metagenomic reads mapped to the Comprehensive Antibiotic Resistance Database (CARD), hereafter referred to as ARG reads, and in the counts of unique ARGs over time, both normalized by the total number of reads in each sample (Figure III.2). Among the GBH treatments applied in Phase I, only the highest concentration increased ARG frequencies over time, either when measured as the number of unique ARGs (GAM F=15.65 p<0.001, Table III.1, Figure III.S2), or as the number of ARG reads (GAM F=15.78 p<0.001, Table III.1, Figure III.S2). The concordance of these two metrics suggests that the effect of GBH on ARGs was not due to a few highly responsive resistance genes, but to multiple unique genes.

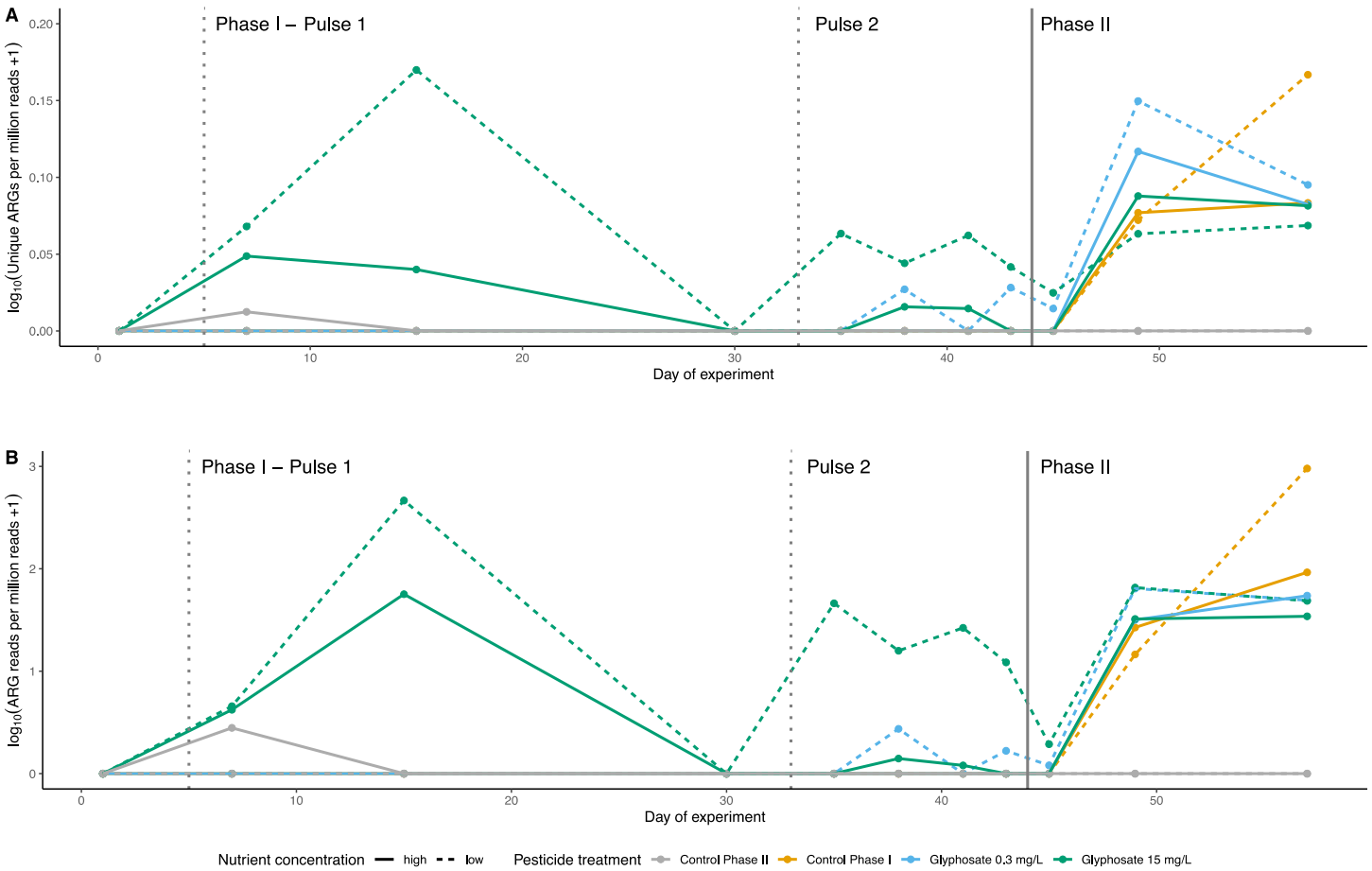


Figure III.2 ARG frequencies increase in GBH treatments over time. (A) Number of unique ARGs per million metagenomic reads and (B) number of metagenomic reads mapped to ARGs per million metagenomic reads vary according to treatment and time. Dashed vertical lines indicate the application of Phase I GBH pulses and solid vertical line the Phase II pulse. The colour code refers to the target glyphosate concentrations in Phase I (pulse 1 and pulse 2), while in Phase II all treated ponds received a target of 40 mg/L glyphosate.

Table III.1 Summary of GAMs showing the effect of GBH on ARG frequencies in phase I only and in both phases[c]

Response variable/ Adjusted R^2	Predictors	Factors[a]	Estimate (SE) or EDF		*t* value or F		*p*-value[b]	
Unique ARG counts per million metagenomic reads ($log_{10}(x+1)$) Adjusted R^2= 65.1% (phase I, n=64)/ 74.6% (both phases, n=88)	*Parametric terms*		Phase I	Both phases	Phase I	Both phases	Phase I	Both phases
	Treatment	Control Phase I	-0.001 (±0.006)	**0.018 (±0.006)**	-0.1	**2.9**	0.890	**0.005***
		Glyphosate 0.3 mg/L	0.003 (±0.006)	**0.023 (±0.006)**	0.5	**3.8**	0.633	**<0.001***
		Glyphosate 15 mg/L	**0.035 (±0.006)**	**0.040 (±0.006)**	**6.2**	**6.7**	**<0.001***	**<0.001***
	Nutrient	High nutrient	**-0.012 (±0.004)**	**-0.011 (±0.004)**	**-3.0**	**-2.7**	**0.005***	**0.009***
	Smooth terms							
	ti(day)	-	1.0	**6.8**	0.02	**2.71**	0.903	**0.010***
	ti(day, by=treatment)	Control Phase I	1.0	**3.2**	0.03	**10.04**	0.861	**<0.001***
		Glyphosate 0.3 mg/L	1.0	**4.7**	0.57	**7.47**	0.453	**<0.001***
		Glyphosate 15 mg/L	**3.9**	**4.3**	**15.65**	**5.01**	**<0.001***	**0.002***
	ti(day, by=nutrient)	High nutrient	1.0	1.0	0.60	0.25	0.444	0.620
ARG reads per million metagenomic reads ($log_{10}(x+1)$) Adjusted R^2= 66.6% (phase I, n=64)/ 77.3% (both phases, n=88)	*Parametric terms*	Factors	Phase I	Both phases	Phase I	Both phases	Phase I	Both phases
	Treatment	Control Phase I	-0.028 (±0.105)	**0.322 (±0.103)**	-0.3	**3.1**	0.790	**0.003***
		Glyphosate 0.3 mg/L	0.013 (±0.105)	**0.320 (±0.103)**	0.1	**3.1**	0.899	**0.003***
		Glyphosate 15 mg/L	**0.678 (±0.105)**	**0.804 (±0.103)**	**6.5**	**7.8**	**<0.001***	**<0.001***
	Nutrient	High nutrient	**-0.197 (±0.074)**	-0.185 (±0.073)	**-2.7**	-2.6	**0.010***	0.013
	Smooth terms							
	ti(day)	-	1.0	**6.7**	0.21	**2.77**	0.648	**0.009***
	ti(day, by=treatment)	Control Phase I	1.0	**3.5**	0.11	**12.27**	0.737	**<0.001***
		Glyphosate 0.3 mg/L	1.0	**2.6**	0.47	**8.93**	0.497	**<0.001***
		Glyphosate 15 mg/L	**3.9**	**4.6**	**15.78**	**6.90**	**<0.001**	**<0.001***
	ti(day, by=nutrient)	High nutrient	1.0	1.0	3.92	1.52	0.053	0.222

[a]When factor is absent, the respective predictor variable is continuous ("day").
[b]Asterisks indicate significant *p*-values after Bonferroni correction (<0.0125).

[c]For each predictor of the model, when it is a parametric term we report the respective parameter estimate with standard errors (SE) and t value; when it is a smooth term, we report the effective degrees of freedom (EDF) and F statistic. Smooths terms are described as the mgcv syntax, and "ti()" phrases are tensor product interactions. *P*-values are reported for each predictor, and reports of significant factors after Bonferroni correction ($p<0.0125$) are highlighted in boldface. A Gaussian residual distribution was used

The first GBH pulse produced a more prominent effect than the second pulse in Phase I (Figure III.2). After each pulse, the ARG frequencies returned close to their baseline, tracking with the community recovery observed at phylum and class taxonomic levels but not at the level of MAGs (Figure III.S1). In Phase II, the single dose of 40 mg/L glyphosate triggered an increase in ARG frequencies across all treated ponds (Figure III.2). ARG frequencies increased over time, due mainly to the Phase II GBH pulse (Table III.1, Figure III.S2). Nutrient enrichment produced a weak but significant effect only when considered alone, not in interaction with time (Table III.1). Overall, these results support the hypothesis that the GBH treatment has the most dominant and strongest positive effect on ARG frequencies over time. The observation that ARG frequencies recover to baseline concordantly with higher taxonomic units but not the finest units (MAGs) suggests that ARGs are encoded by a wide range of distantly related bacteria and the number of species containing ARGs may be large, rather than restricted to a few closely related strains.

GBH selects for specific gene functions, including antibiotic efflux

To assess how GBH affected known gene functions beyond ARGs in the bacterial communities, we built PRCs based on SEED annotations of genes in the metagenomes. The PRCs revealed a clear effect of GBH on the composition of gene functions (Figure III.S3), similar to the treatment effect detected in the PRCs for community composition at the phylum and class levels (Figure III.S1B-C). In Phase I, the first pulse of 15 mg/L glyphosate induced greater deviations from controls than the second pulse. In Phase II, all ponds receiving 40 mg/L glyphosate deviated from the controls. Resistance to antibiotics is among the functions positively affected by GBH treatment, as indicated by the positive scores of the SEED subsystems "Virulence, Disease and Defense", at level 1 (Figure III.S3A), and "Resistance to antibiotics and toxic compounds", at level 2 (Figure III.S3B). Table III.S1 shows the complete list of PRC scores for all SEED subsystems at

levels 1 and 2. Membrane transport (level 1, Figure III.S3A), such as the ATP-binding cassette (ABC) transporters (level 2, Figure III.S3B), are among the positively selected functions. These genes could plausibly change cell permeability to various molecules, including glyphosate.

To assess the effects of GBH on ARGs at a higher level of resolution, we built another set of PRCs based on ARG profiles predicted from reads mapping to CARD. The resulting PRC plot showed a prominent effect of the first and second pulses of 15 mg/L of glyphosate in Phase I (Figure III.3). In Phase II, the GBH had an effect in all treatments that received a last pulse (40 mg/L glyphosate). This result is consistent with the greater effect of the large Phase II pulse compared to smaller Phase I pulses on total ARG frequencies (Figure III.2 and Figure III.S2). The two principal resistance mechanisms of the ARGs annotated by CARD are antibiotic efflux and antibiotic inactivation (shown respectively in blue and red text in Figure III.3). Genes encoding antibiotic efflux functions were more often found with positive PRCs scores (Fisher's exact test, p=0.013), suggesting that they tend to be selected more often than other ARGs in the presence of GBH. This result supports the hypothesis that membrane transporters used for antibiotic efflux could also play a role in exporting glyphosate from bacterial cells.

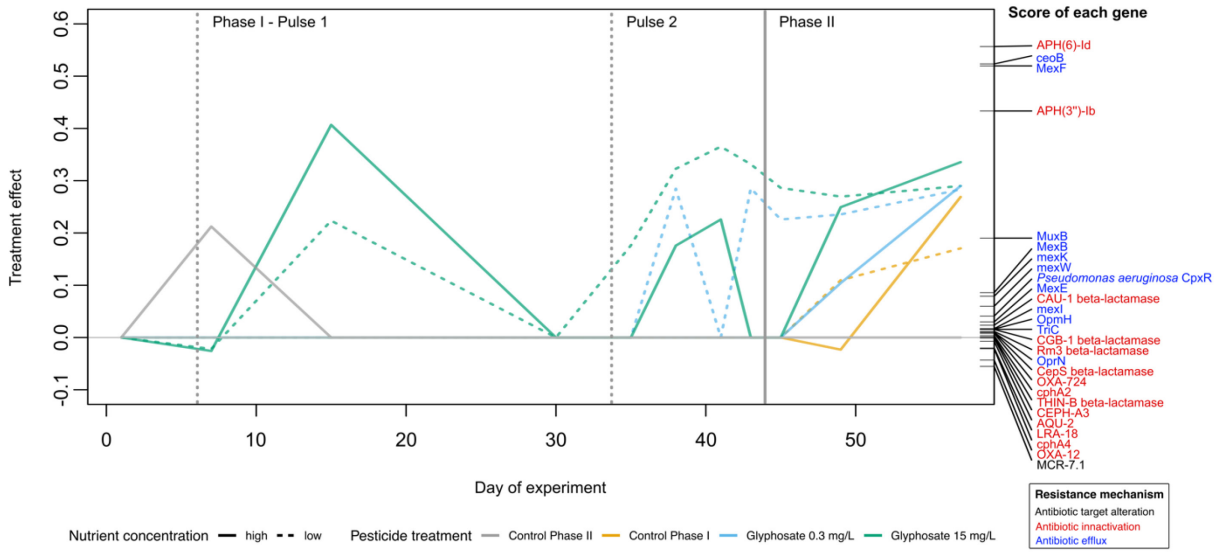


Figure III.3 GBH skews composition of ARGs in favor of antibiotic efflux pumps. Principal Response Curves (PRCs) illustrating divergence (relative to controls) in the composition of ARGs in response to GBH exposure. The left y-axis represents the

magnitude or ARG compositional response, while the right y-axis represents individual gene scores (i.e., relative contribution to overall compositional changes). Gene names (ARO) are colour-coded based on their mechanism of resistance. Dashed vertical lines indicate the timing of GBH pulses in Phase I, and the solid vertical line represents the pulse in Phase II. The zero line (y=0) represents the low nutrient control pond from both Phase I and II. The PRC explains 30% of the total variance (PERMUTEST, F=22.8, p=0.024). Treatments and time interactively explain 74.8% of the variance while 25% is explained by time alone.

Connecting resistance genes to genomes and plasmids

Thus far, our results have only considered ARGs outside the context of the bacterial genomes or plasmids in which they occur. On average, 71% (± 3; range = 45–94%, Table III.S2) of ARG reads (those mapping to CARD) across all samples also mapped to MAGs, implying that MAGs captured a large fraction of ARG reads in the metagenomes. We identified putative plasmids in 390 MAGs, with an average of 43 plasmid contigs per MAG (min=1, max=520, SE=3.5, Table III.S3). However, only 27 plasmid contigs in total were annotated with ARGs. Out of a total of 188 MAGs with ARGs, only 24 (13%) of them had at least one ARG identified in a potential plasmid. Although some ARGs are clearly encoded on plasmids, ARGs are more frequently associated with genomes than with MAG plasmids in our study.

Of the 426 total MAGs, only 20 recruited 100 or more ARG reads, and the classification of their EPSPS genes varied (Figure III.S4, S5). To visualize which ARGs were more abundant in GBH treatments and in which MAGs they were found, we examined the frequency of metagenomic reads mapped to ARGs according to their antibiotic resistance ontology (ARO) classification (top graphs in Figure III.S4 and Figure III.S5) as well as the proportion of these reads that were mapped to MAGs (bottom graphs in Figure III.S3 and Figure III.S5). These visualizations confirmed that efflux pumps (e.g. *mex* genes) increased in frequency in response to GBH. The relative abundance of *mex* genes is strongly associated with a *Pseudomonas putida* MAG (Figure III.S4; bottom right panel) but is sometimes also associated with other MAGs such as *Aeromonas veronii* (Figure III.S4), Oxalobacteraceae, and *Azospirillum* (Figure III.S5). It is thus likely that GBH selects for efflux pump genes in multiple different genomic backgrounds.

The number of ARGs encoded in a MAG predicts its frequency after severe GBH exposure

Collectively, our results suggest an important role for ARGs, and efflux pumps in particular, in allowing bacterioplankton to survive and grow in the presence of a GBH. We next asked, what is the importance of ARGs relative to genetic variation in the glyphosate target enzyme, EPSPS? Based on known sequence variation in the EPSPS encoding gene, we were able to classify MAGs as putatively glyphosate resistant, sensitive, or unclassified. We also defined a MAG's antibiotic resistance potential as the number of ARGs they contain (i.*e.* the number of RGI strict hits). These definitions are genetic predictions based on the presence or absence of resistance-associated genes in MAGs, but they do not represent confirmed phenotypes. We nonetheless tested the extent to which these genomic features were predictive of a MAG's average relative abundance across ponds at the end of the experiment, after receiving 40 mg/L glyphosate in Phase II. We found that MAGs encoding more unique ARGs tended to have higher relative abundance after receiving the GBH pulse in Phase II (Figure III.4A, Table III.2). This effect of antibiotic resistance potential was highly significant (multiple linear regression model, t=9.53 p<0.001, Table III.2), and was not observed in control ponds that did not receive the Phase II pulse (Figure III.S6A; t=2.26 p=0.025; not significant after Bonferroni correction, Table III.S1). The relative abundance of MAGs at the end of the experiment in these control ponds was predicted by their relative abundance in Phase I (40% of variance explained; Table III.S4), consistently with temporal autocorrelation (*e.g.* due to fluctuations in species abundances unrelated to experimental treatments). In contrast to the strong effect of ARGs on predicting MAG relative abundance after GBH stress (17% of variance explained; Table III.2), EPSPS classification explained only 2% of the variation – in both Phase II treatment and control ponds.

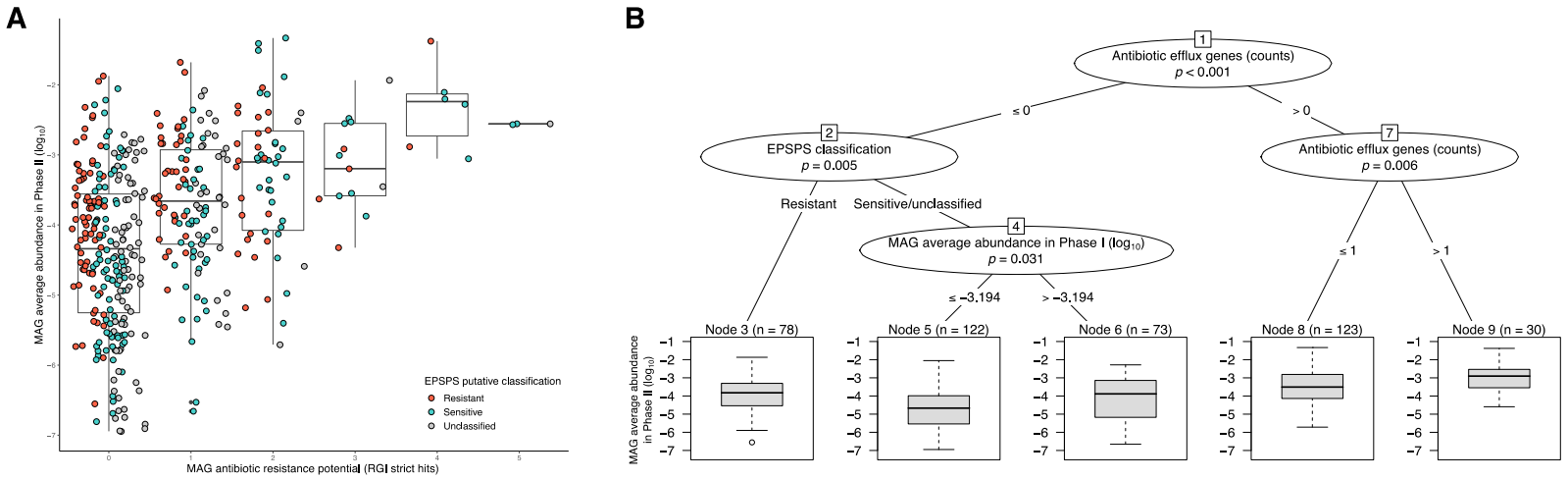


Figure III.4 Antibiotic resistance potential predicts MAG relative abundance after severe GBH stress. (A) Boxplots show a positive correlation between MAGs abundance in Phase II and their potential for antibiotic resistance. Each dot represents a MAG that is color-coded based on the predicted resistance of their EPSPS. A slight offset on x-axis (jitter) was introduced to facilitate data visualization. See Table III.2 for regression coefficients. (B) Regression tree confirms the significance of the correlation seen in (A), particularly for antibiotic efflux genes. Two other factors were also included, and have small effects on MAG relative abundance in Phase II: the EPSPS classification and the average abundance of MAGs in Phase I.

Table III.2 Multiple linear regression model and variance partitioning of MAGs abundance in phase II in treatment mesocosms[a]

Predictor	Estimate (SE)	*t* value	*p*-value	Explained variance (%)
EPSPS classification:				2
- Sensitive	0.002 (±0.127)	0.02	0.987	
- **Resistant**	**0.413 (±0.133)**	**3.11**	**0.002***	
MAG antibiotic resistance potential	**0.496 (±0.052)**	**9.53**	**<0.001***	17
MAG mean abundance in Phase I treatment mesocosms (log_{10})	**0.178 (±0.066)**	**2.69**	**0.007**	1; Residuals: 79

[a]*P*-values are reported for each predictor, asterisks indicate significant *p*-values after Bonferroni correction ($p<0.0125$) and reports of significant factors are highlighted in bold. Adjusted R-squared equals 21.1% for MAG persistence in treatments ($n=426$, F-statistic: 29.5).

To further explore these results, we used a regression tree analysis to identify the drivers of MAG abundance at the end of Phase II (Figure III.4B). Instead of combining the three major classes of ARGs (antibiotic target alteration, antibiotic inactivation and antibiotic efflux), we used each as a separate predictor in the regression tree. The first division splits MAGs with at least one antibiotic efflux gene (Figure III.4B, node 7) which were on average more abundant post-GBH pulse than those without efflux genes (Figure III.4B, node 2). Among MAGs with efflux genes, the more genes they had, the higher their abundance. Among MAGs without antibiotic efflux genes, the EPSPS classification was an important driver of their abundance, followed by the MAG's average abundance in Phase I. In the absence of a GBH pulse in Phase II, the primary driver of MAG abundance in Phase II controls was their mean relative abundance in Phase I (Figure III.S6). Control pond regression trees also included a split between resistant/sensitive and unclassified EPSPS, which is difficult to interpret biologically and likely attributable to noise. This could also explain why 2% of the variation in MAG relative abundance in control ponds was explained by EPSPS class. Together, these results indicate that a bacterial genome's ARG coding potential is predictive of its ability to persist in the face of GBH stress – more so than the class of EPSPS enzyme it encodes.

Discussion

Our mesocosm experiment used deep metagenomic sequencing to assess the effect of a GBH on microbial genes and genomes in semi-natural freshwater bacterial communities. We show that exposure to Roundup® at high concentrations (15 mg/L and 40 mg/L glyphosate) changes community composition and increases the frequency of ARGs in freshwater bacterioplankton. Consistent with our previous 16S rRNA gene amplicon sequencing from the same experiment (Barbosa da Costa et al., 2021), we found that more inclusive taxonomic groupings were resilient to GBH pulses, whereas the precise MAG composition never recovered. Moreover, we show that the abundance of MAGs after severe contamination (40 mg/L glyphosate) was predictable based on the

number of ARGs encoded, and such 'successful' MAGs tended to have at least one antibiotic efflux gene annotated in their genome. The effect of GBH on ARGs is likely due to cross-resistance, since the multidrug efflux pumps which rise in frequency in response to GBH could potentially transport glyphosate in addition to antibiotics [18]. Alternatively, co-resistance could play a role if GBH selects for glyphosate-resistant bacterial genomes or genetic elements (rather than specific genes) that happen also to encode ARGs. While we cannot exclude a role for co-resistance entirely, the cross-resistance model is more plausible since efflux genes are strongly affected, likely in multiple independent genomic backgrounds. As discussed in detail below, direct selection on the EPSPS locus appears to be weak, implying that ARGs are unlikely to achieve high frequency due to genetic linkage with resistant EPSPS alleles.

There are several explanations for the observation that ARG frequencies and higher taxonomic units – phyla and classes but not MAGs – concurrently increase then recover to baseline following GBH pulses. One technical explanation is that ARG families are defined at levels of genetic similarity more in line with phyla or classes than with finer taxonomic levels. It is also possible that the recovery of ARG frequencies to baseline after Phase I pulses could be due to gene loss and horizontal transfer events (as discussed further below). However, even if such events occur, they are not sufficient to obscure the predictability of ‘successful’ MAGs after a severe GBH pulse – which require relatively tight linkage of ARGs with genomes.

An association between glyphosate, ARGs, and mobile genetic elements has been previously found in soil microbiomes, as demonstrated in a recent study combining experimental microcosms and environmental data from agricultural field sites in China (Liao et al., 2021). Through laboratory assays of three bacterial strains, the authors quantified the conjugation frequency of a multidrug resistance plasmid induced by glyphosate and further investigated changes in cell membrane permeability. They detected a significant increase in conjugation frequency and augmented cell membrane permeability in the presence of glyphosate, suggesting that glyphosate stress increases membrane permeability, thereby promoting plasmid movement. Here, we provide additional support for the hypothesis that cell membrane permeability is altered in the presence of a GBH, as demonstrated by the selection of membrane transport

mechanisms, such as ABC transporters (Davidson & Chen, 2004) among the annotated gene functions most responsive to our GBH treatments. In contrast, although we did not quantify the frequency of conjugation in our experiment, we did identify some ARGs located on putative plasmids which were present in a small fraction of the bacterial community. Of the MAGs encoding ARGs, only 13% contained a predicted plasmid-encoded ARG. It is possible that unassembled plasmids or plasmids not associated with MAGs could harbor ARGs. Including such plasmids would not be expected to change our main conclusion that ARGs are more predictive of MAG frequency post-GBH exposure than EPSPS. In addition to plasmids, other mechanisms also contribute to horizontal gene transfer between bacteria, such as phage-mediated transduction and transformation (Wiedenbeck & Cohan, 2011), and future studies could test how these processes may be affected by GBH stress.

Strikingly, antibiotic resistance potential, particularly the presence of antibiotic efflux genes, was more important than the EPSPS classification in explaining variation in MAG abundance in Phase II, after a high GBH pulse. This evidence of cross-resistance in semi-natural communities may help explain why, in previous experiments also performed with complex communities, bacterial strains with the sensitive EPSPS encoding gene were resistant to glyphosate, as was the case in two strains of *Snodgrassella alvi* in the bee gut microbiome (Motta et al., 2018). Although EPSPS alleles were weakly predictive of MAG relative abundance after the phase II GBH pulse, their effects were clearly secondary to the strong effects of ARGs. Computational gene annotations of both ARGs and resistant or sensitive EPSPS have limitations because they are based on sequence similarity, not on phenotypic measurements. Therefore, we cannot exclude a role for EPSPS alleles in conferring GBH resistance in nature, but their effects were small in our experiment. Together, our results strongly suggest that ARGs (and efflux pumps in particular) could be more relevant to glyphosate resistance in nature than mutations in the glyphosate target enzyme.

Our study aligns with previous single-strain laboratory evidence that antibiotic resistance may enhance bacterial survival in the presence of pesticides. Laboratory assays of bacterial isolates showed that exposure to agrochemicals accelerated the rate of antibiotic resistance evolution (Kurenbach et al., 2018; Xing et al., 2020). In other

studies, depending on the combination of herbicide and antibiotics tested, herbicides increased or decreased antibiotic susceptibility of bacterial strains (Kurenbach et al., 2017, 2015). In our experiment, GBH pulses caused a general increase in ARG frequency in the bacterioplankton community, suggesting an overall positive effect of GBH on resistance to antibiotics. This does not exclude the possibility that GBH could also select against specific ARGs or mutations, making bacteria less resistant to certain antibiotics. As a metagenomic study, we only measured genetic correlates of resistance and future experiments will be needed to examine resistance phenotypes.

Our study shows that antibiotic efflux is a major mechanism of antibiotic resistance that is cross-selected by GBH stress, which also corroborates previous laboratory assays. It has been shown that the targeted deletion of efflux pump genes can neutralize the increased tolerance to kanamycin and ciprofloxacin in *Escherichia coli* and *Salmonella enterica* serovar Typhimurium in the presence of GBH (Kurenbach et al., 2017, 2015). Here we provide evidence that efflux pumps may also provide resistance to both glyphosate and certain antibiotics in a more natural and complex system. Whether all efflux pumps are equally capable of transporting various molecules out of the cell remains to be seen, and other resistance mechanisms could also play a role. Resistance to both antibiotics and GBH could also be modulated by changing the expression of efflux pumps. While we used a metagenomic approach to show how GBH affects ARG frequencies in a community, this does not exclude the possibility of changes in gene expression which could be tracked using metatranscriptomics in future experiments.

It should be noted that we used a commercial Roundup® formulation of the herbicide glyphosate, which includes other constituents that may also influence microbial communities and cellular physiology. For example, the surfactant polyethoxylenamine (POEA) has produced negative effects on *Vibrio fischeri* at lower concentrations than pure glyphosate acid (Tsui & Chu, 2003). However, given that our results are in general agreement with previous soil experiments using pure glyphosate (Liao et al., 2021), we believe that our findings are at least in part attributable to an effect of glyphosate itself. Furthermore, regardless of whether it is glyphosate or other constituents of GBH that drive cross-selection of ARGs, assessing the risks associated with commercial formulations is

ecologically relevant, as these formulations are used in agriculture fields and lawns (Mesnage & Antoniou, 2018).

On an applied level, the safety assessment process for pesticides such as glyphosate, currently based on toxicity to model eukaryotic organisms (CCME, 2012; EPA, 2019), could also consider the potential effects on bacterioplankton and ARGs. Our results highlight the role of GBH contamination as an indirect selective pressure favouring ARGs in natural communities. Although glyphosate concentrations as high as the ones inducing this effect (i.e. 15 mg/L and 40 mg/L) are rarely found in nature, there are reports of even higher glyphosate levels during the rainy season close to agricultural fields, as observed in Argentina (105 mg/L) for example (Brovini et al., 2021). Additionally, currently regulated acceptable concentrations of glyphosate in freshwaters in the USA and Canada for short-term exposure (1-4 days) are close to the concentrations used in our experiment (respectively 49.9 mg/L (EPA, 2019) and 27 mg/L (CCME, 2012)). Here we have shown that ARG frequencies can rise dramatically just a few days after GBH treatment at such doses. The extent to which these ARGs, and the bacteria that encode them, can be mobilized across aquatic ecosystems, and from these ecosystems into animals and humans, remains to be seen.

Data accessibility: Sequence data of the 88 metagenomic samples were submitted to NCBI SRA (BioProject PRJNA767443, accession numbers SRR16126824-SRR16126911) and the genomes of 426 predicted MAGs have been deposited and associated to the same BioProject (BioSample accession numbers in Table III.S3).

Acknowledgements: We are grateful to D. Maneli, C. Normandin, A. Arkilanian and T. Jagadeesh for their assistance in the field, to J. Marleau and O.M. Pérez-Carrascal for their assistance in the laboratory and to O.M. Pérez-Carrascal for the advice on bioinformatic analyses. We also thank the two anonymous reviewers for their feedback on a previous version of this manuscript.

Funding information: This study was supported by a Canada Research Chair and NSERC Discovery Grant to B.J.S. N.B.C. was funded by FRQNT and NSERC-

CREATE/GRIL fellowships. M-P.H. was funded by NSERC and NSERC-CREATE/GRIL. V.F. was supported by an NSERC postdoctoral fellowship. LEAP was built and operated with funds from a CFI Leaders Opportunity Fund, NSERC Discovery Grant and the Liber Ero Chair to A.G.

Supplementary information

Supplementary figures

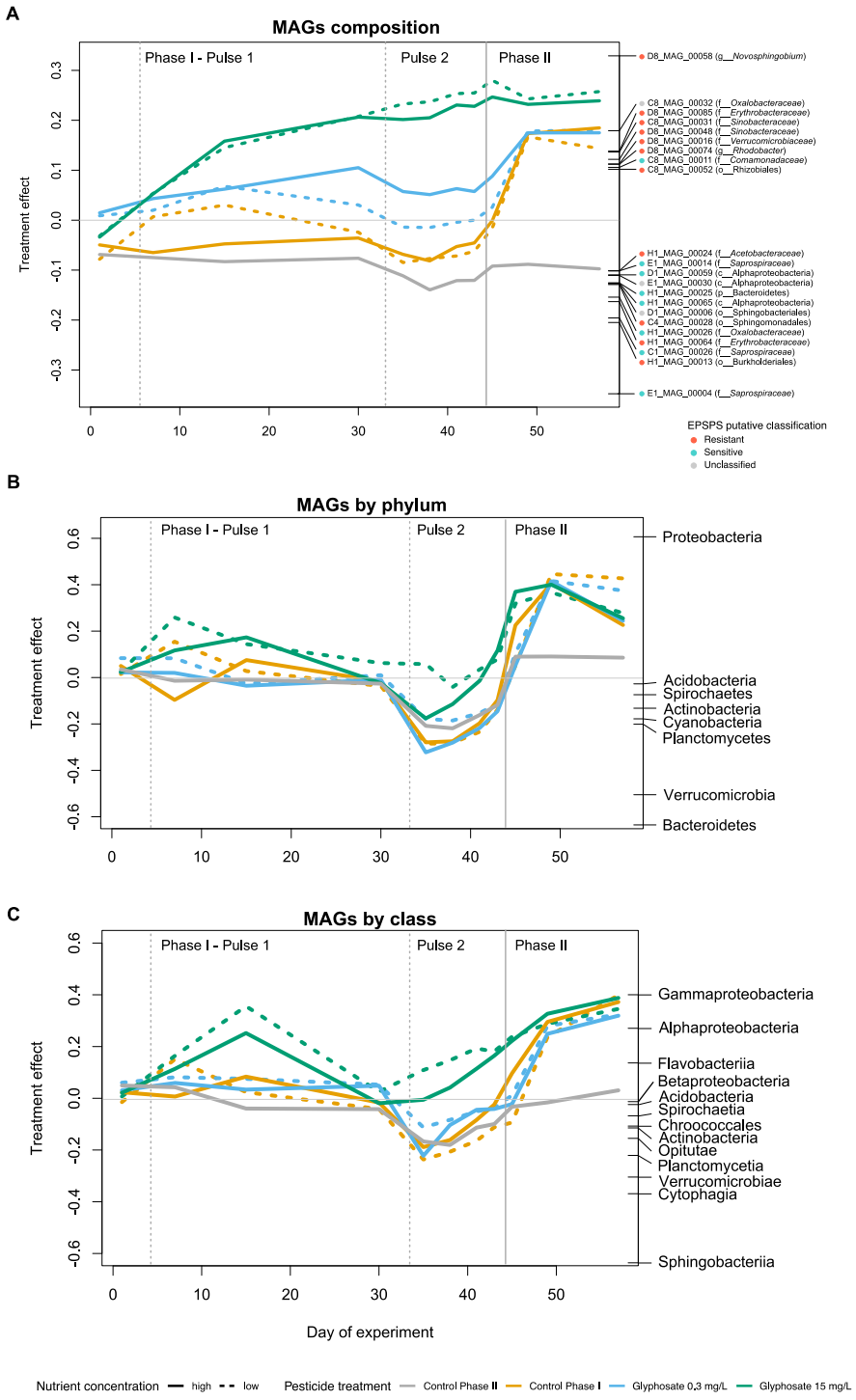


Figure III.S1 Principal Response Curves (PRCs) of the experimental treatment effect on the composition of MAGs (A) and MAGs grouped at the phylum (B) and class (C) level.

Treatment effect is shown in the left y-axis while taxa scores (proportional to their contribution to the treatment effect) are shown in the right y-axis. Dashed vertical lines indicate the application of Phase I glyphosate pulses and solid vertical line the Phase II glyphosate pulse. Glyphosate concentration of pulses applied in Phase I (dose 1 and dose 2) are indicated by the legend, while in Phase II all treatments received 40 mg/L of glyphosate, except the Phase II controls. Treatment effect zero is equivalent to the low nutrient control Phase II pond. (A) Only taxa score higher than 0.1 are shown, shown axis explained 16.1% of total variance, PERMUTEST $F=14.6$ $p=0.011$, (B) 51.5% of total variance explained, PERMUTEST $F=44.4$ $p=0.024$, (C) 31.8% of total variance explained, PERMUTEST $F=34.3$ $p=0.008$).

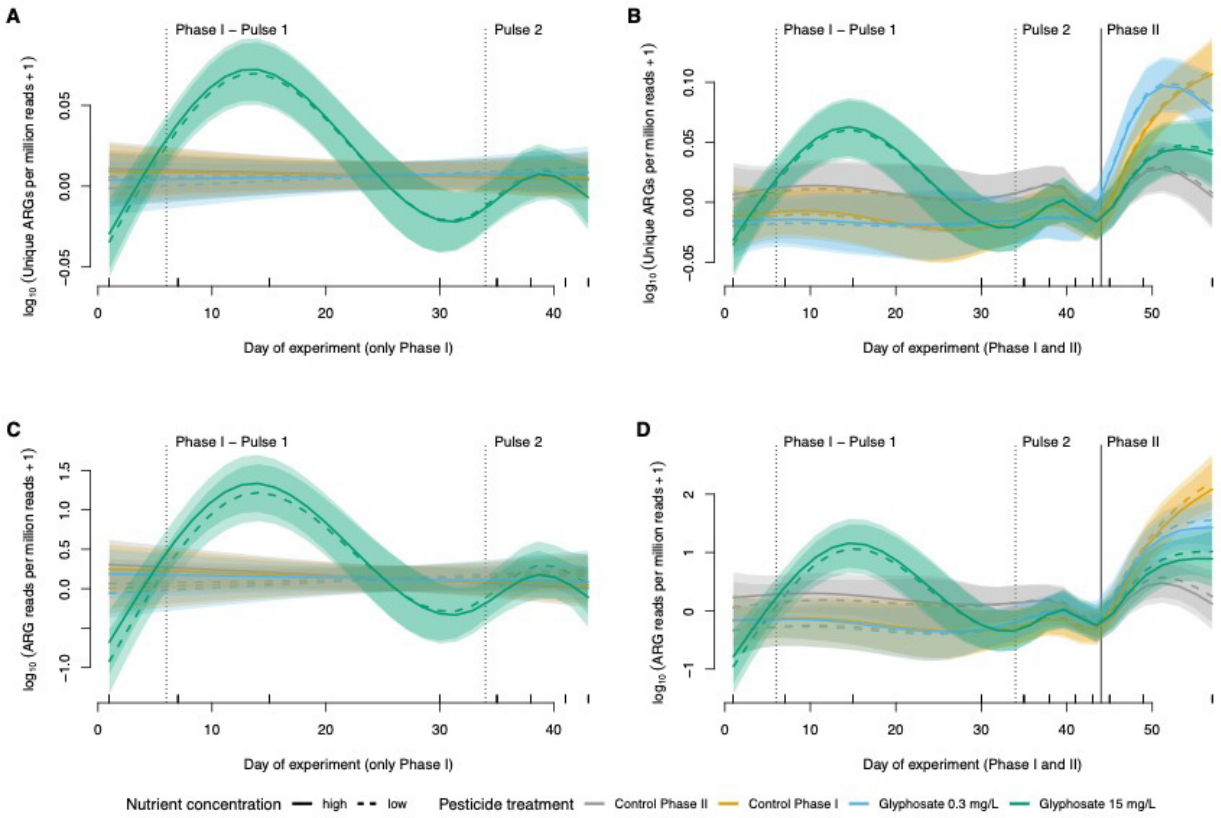


Figure III.S2 Glyphosate increases ARG frequencies in experimental ponds. GAMs illustrating the time-dependent effect of GBH and nutrient treatments on unique ARGs in Phase I (A), in both Phase I and II (B), on ARG reads in Phase I (C), in both Phase I and Phase II (D). Dashed vertical lines indicate the application of Phase I GBH pulses and solid vertical line the Phase II pulse. Glyphosate acid concentration of pulses applied in Phase I (dose 1 and dose 2) are indicated in the legend, while in Phase II, all treatments received 40 mg/L, except the Control Phase II. Shaded bars indicate a confidence interval of 95%.

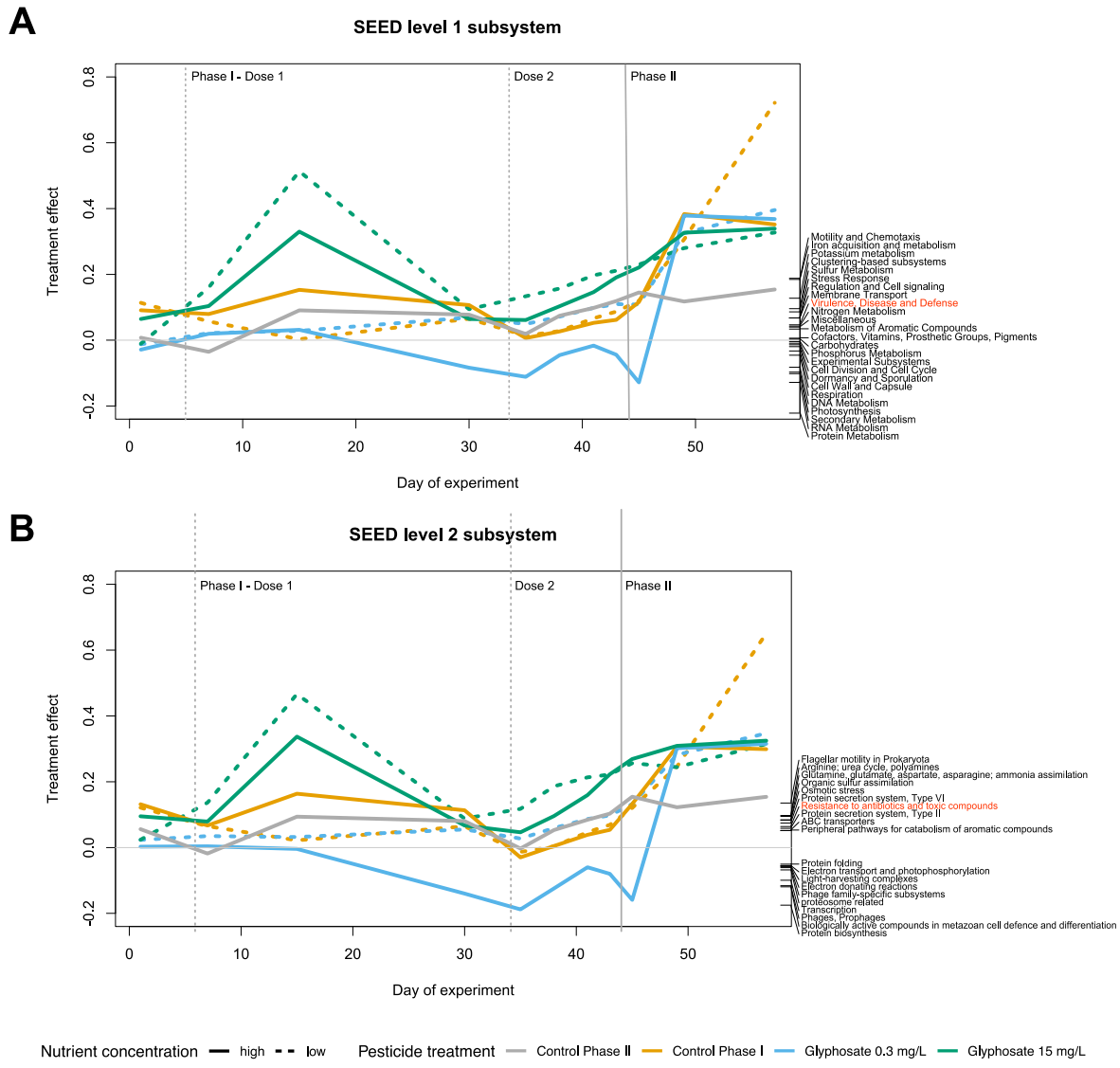


Figure III.S3 PRCs of the experimental treatment effect on the composition of gene functional profiles predicted from metagenomic reads grouped according to (A) SEED subsystem level 1 and (B) level 2. Treatment effect is shown in the left y-axis while scores of genes (proportional to their contribution to the treatment effect) are shown in the right y-axis. Dashed vertical lines indicate the application of Phase I glyphosate pulses and solid vertical line the Phase II glyphosate pulse. Glyphosate concentration of pulses applied in Phase I (dose 1 and dose 2) are indicated by the legend, while in Phase II all treatments received 40 mg/L of glyphosate, except the Phase II controls. Treatment effect zero is equivalent to the low nutrient control Phase II pond. Function of resistance to antibiotics is highlighted in red according to how it is named in (A) SEED subsystem level 1 (50.9% of total variance explained, PERMUTEST F=43.1 p=0.023) and (B) SEED subsystem level 2 (33.1% of total variance explained, PERMUTEST F=25.8 p=0.027), where only scores with absolute values larger than 0.05 are reported (all scores are shown in Table III.S1).

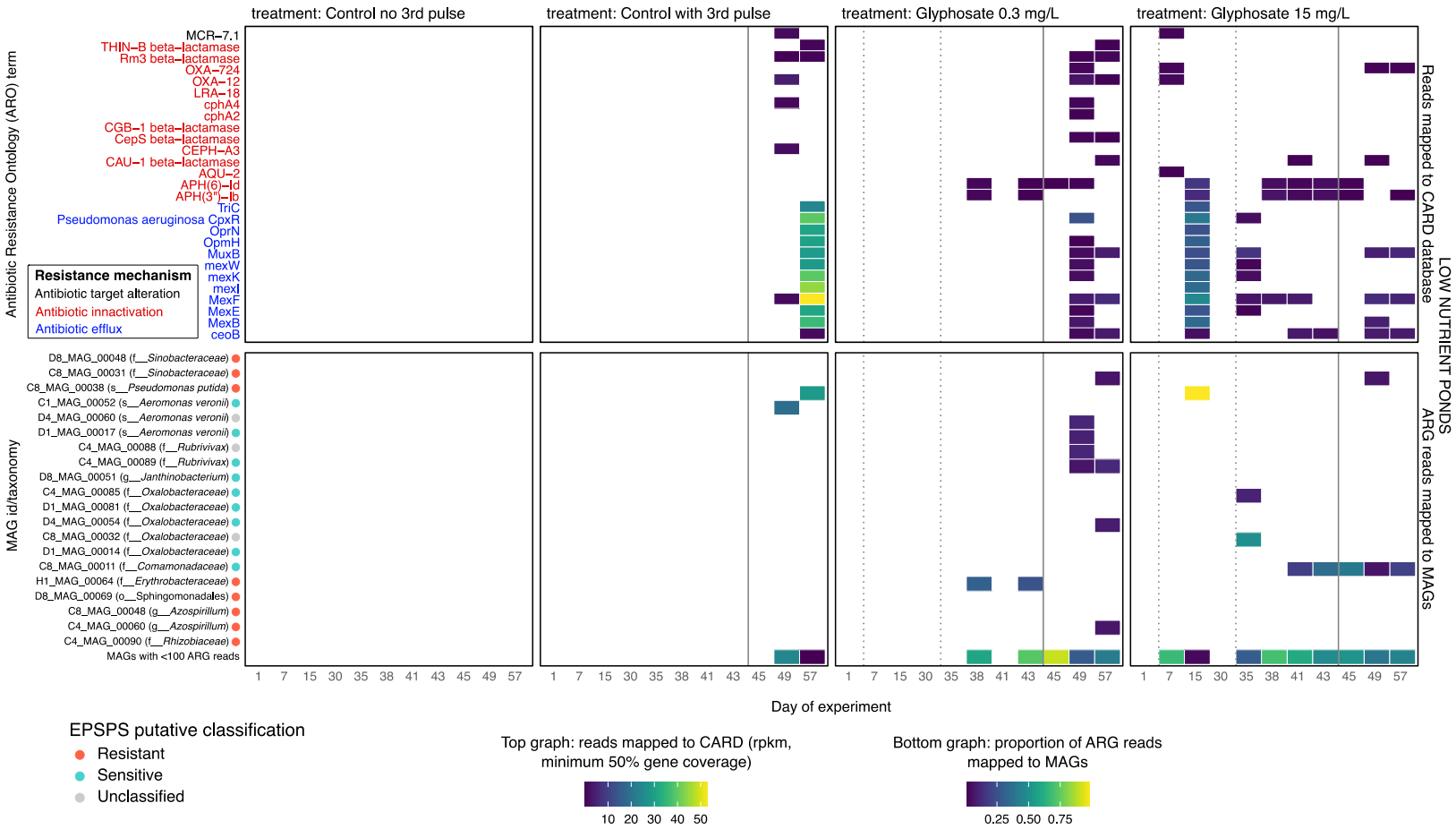


Figure III.S4 Metagenomic reads mapped to ARGs classified according to their ARO (top graph) and ARG reads mapped to MAGs (bottom graph) in low nutrient ponds. MAG identities are followed by their finest taxonomic assignment (o=order, f=family, g=genus, s=species). Only alignments with MAPQ>10 were tallied. Dashed vertical lines represent Phase I GBH and solid vertical lines are Phase II pulses (all at 40 mg/L glyphosate).

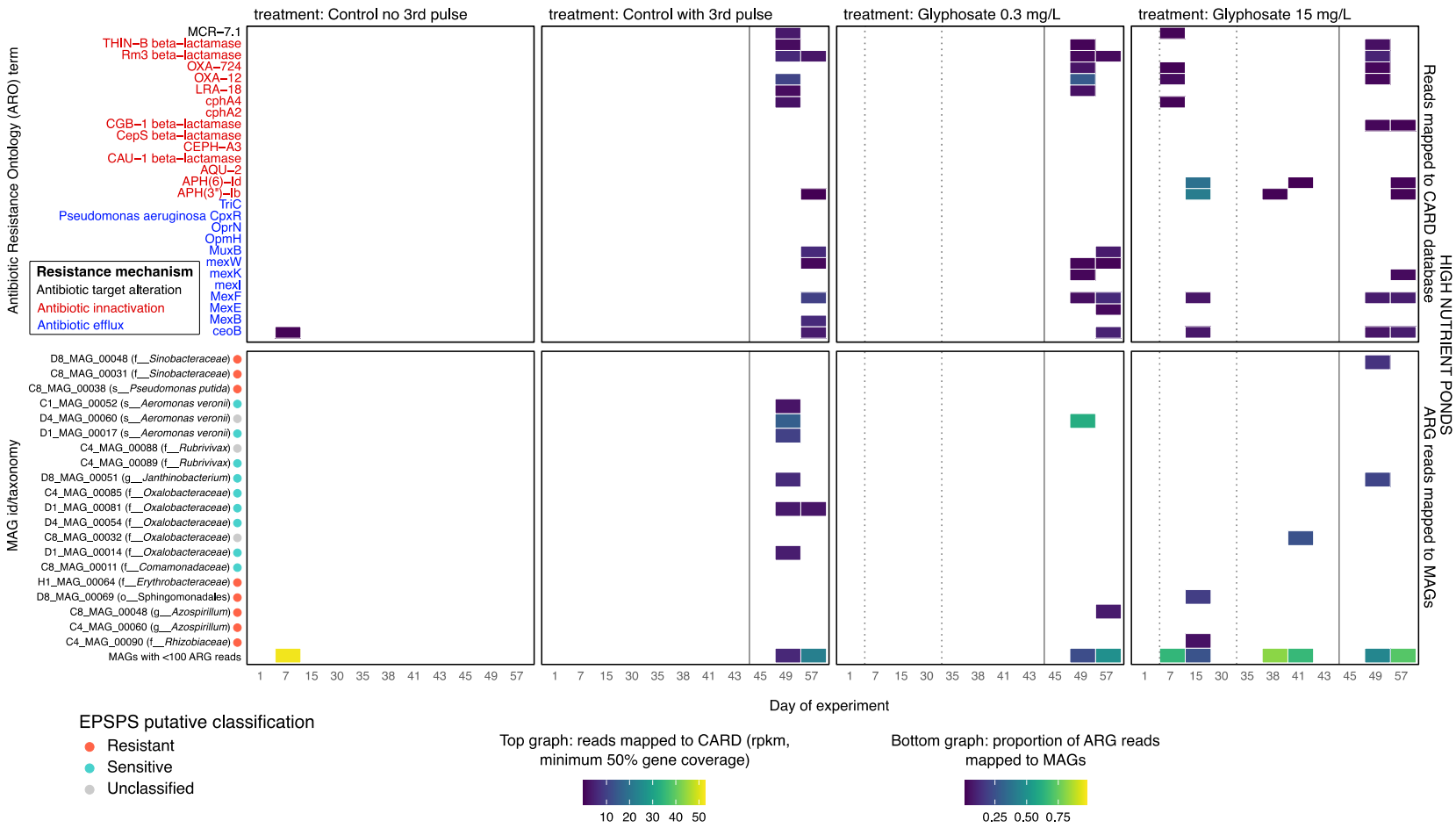


Figure III.S 5 Metagenomic reads mapped to ARGs classified according to their ARO (top graph) and ARG reads mapped to MAGs (bottom graph) in high nutrient ponds. MAG identities are followed by their finest taxonomic assignment (o=order, f=family, g=genus, s=species). Only alignments with MAPQ>10 were tallied. Dashed vertical lines represent Phase I GBH pulses and solid vertical lines are Phase II pulses (all at 40 mg/L glyphosate).

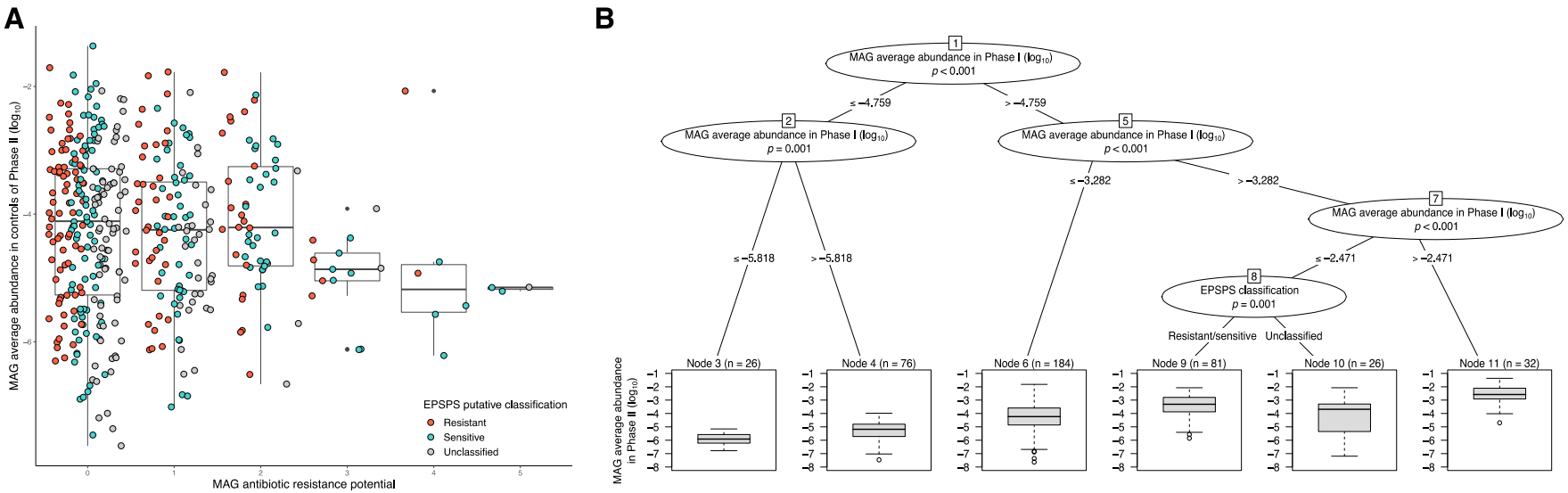


Figure III.S6 MAG mean relative abundance in controls of Phase II as a function of antibiotic resistance potential (or the amount of ARGs annotated to their genomes) and the classification of EPSPS enzyme (resistant, sensitive or unclassified). (A) Series of boxplots show the absence of correlation between MAGs abundance in Phase II and their potential for antibiotic resistance. Each dot represents a MAG that is color-coded according to the potential resistance of their EPSPS. To facilitate visualization, a small amount of random variation (jitter) was added so dots would not overlap. Table III.2 reports statistics of a linear model that tested how MAG abundance in Phase II controls could be explained by EPSPS classification, antibiotic resistance potential and MAG abundance in Phase I. (B) Regression tree with MAG abundance in controls of Phase II as the response variable and the following predictors: the EPSPS enzyme classification, the number of ARGs classified as antibiotic efflux, antibiotic inactivation or target alteration, and the MAG relative abundance in Phase I.

Supplementary tables

Table III.S1 PRC scores from functional annotations shown in Figure III.S3

SEED subsystem classification	PRC score	SEED subsystem Level
Motility and Chemotaxis	0.1882	1
Iron acquisition and metabolism	0.1849	1
Potassium metabolism	0.1276	1
Clustering-based subsystems	0.0958	1
Sulfur Metabolism	0.0862	1
Stress Response	0.0678	1
Regulation and Cell signaling	0.0677	1
Membrane Transport	0.0676	1
Virulence, Disease and Defense	0.0467	1
Nitrogen Metabolism	0.0426	1
Miscellaneous	0.0400	1
Metabolism of Aromatic Compounds	0.0339	1
Cofactors, Vitamins, Prosthetic Groups, Pigments	0.0061	1
Carbohydrates	0.0035	1
Phosphorus Metabolism	-0.0040	1
Experimental Subsystems	-0.0092	1
Cell Division and Cell Cycle	-0.0139	1
Dormancy and Sporulation	-0.0194	1
Cell Wall and Capsule	-0.0329	1
Respiration	-0.0457	1
DNA Metabolism	-0.0821	1
Photosynthesis	-0.0972	1
Secondary Metabolism	-0.1015	1
RNA Metabolism	-0.1281	1
Protein Metabolism	-0.2217	1
Flagellar motility in Prokaryota	0.1353	2
Arginine; urea cycle, polyamines	0.0976	2
Glutamine, glutamate, aspartate, asparagine; ammonia assimilation	0.0946	2
Organic sulfur assimilation	0.0843	2
Osmotic stress	0.0827	2
Protein secretion system, Type VI	0.0744	2
Resistance to antibiotics and toxic compounds	0.0640	2
Protein secretion system, Type II	0.0584	2
ABC transporters	0.0520	2
Peripheral pathways for catabolism of aromatic compounds	0.0516	2

Methylamine utilization	0.0489	2
Central carbohydrate metabolism	0.0455	2
Sugar alcohols	0.0454	2
May be related to ADP-phosphoribose and NAD-dependent acetylation	0.0452	2
Clustering-based subsystems	0.0430	2
Putative GGDEF domain protein related to agglutinin secretion	0.0427	2
Putative asociate of RNA polymerase sigma-54 factor rpoN	0.0424	2
Quinone cofactors	0.0404	2
alpha-proteobacterial cluster of hypotheticals	0.0404	2
Tricarboxilate (malonate, propionate?) transport	0.0398	2
Triacylglycerols	0.0385	2
Proline and 4-hydroxyproline	0.0371	2
Aminosugars	0.0364	2
Fermentation	0.0357	2
Metabolism of central aromatic intermediates	0.0338	2
Protein secretion system, Type VIII (Extracellular nucleation/precipitation pathway, ENP)	0.0333	2
Detoxification	0.0326	2
Lysine, threonine, methionine, and cysteine	0.0306	2
Prophage	0.0292	2
Choline bitartrate degradation, putative	0.0291	2
Tetrapyrroles	0.0282	2
Protein secretion system, Type III	0.0267	2
TldD cluster	0.0254	2
Electron accepting reactions	0.0231	2
Gene Transfer Agent (GTA)	0.0211	2
General secretion system/Phosphate-binding DING proteins cluster	0.0209	2
Phosphate metabolism	0.0204	2
Bacterial cytostatics, differentiation factors and antibiotics	0.0187	2
Transposable elements	0.0176	2
Translation	0.0174	2
May be related to amine metabolism	0.0165	2
Carbohydrates	0.0164	2
Social motility and nonflagellar swimming in bacteria	0.0159	2
Programmed Cell Death and Toxin-antitoxin Systems	0.0145	2
Adhesion	0.0143	2

Flagella protein?	0.0143	2
Hypothetical aromatic compound degradation cluster	0.0142	2
Chemotaxis, response regulators	0.0141	2
Selenoproteins	0.0138	2
Purines	0.0118	2
heat shock, cell division, proteases, and a methyltransferase	0.0117	2
NAD and NADP	0.0102	2
Putrescine/GABA utilization cluster-temporal,to add to SSs	0.0101	2
Ubiquinol-cytochrome C chaperone locus	0.0101	2
hypthetical clustered with tRNA modification	0.0095	2
Alanine, serine, and glycine	0.0088	2
Two related proteases	0.0088	2
Hypothetical in Lysine biosynthetic cluster	0.0080	2
Chromosome Replication	0.0080	2
ATP synthases	0.0078	2
Pyruvate kinase associated cluster	0.0078	2
Protein and nucleoprotein secretion system, Type IV	0.0078	2
Denitrification	0.0073	2
methane monooxygenase cluster (? can be phenol-monoox in some organisms?)	0.0071	2
Drug resistance or antibiotic biosynthesis related cluster	0.0066	2
Probably Ybbk-related hypothetical membrane proteins	0.0064	2
Protein secretion system, Type VII (Chaperone/Usher pathway, CU)	0.0055	2
Permease proteins, subunit of DNA pol, and the Val-tRNA synthetase	0.0046	2
Ribosome-related cluster	0.0038	2
Shiga toxin cluster	0.0037	2
A cluster-based salvage pathway	0.0034	2
DNA uptake, competence	0.0024	2
Gram-Negative cell wall components	0.0022	2
DNA polymerase III epsilon cluster	0.0022	2
Toxins and superantigens	0.0009	2
CRISPs	0.0006	2
Regulation of virulence	0.0005	2
Periplasmic Stress	0.0004	2
DNA metabolism	0.0003	2
Magnetotaxis	0.0001	2

DNA recombination	0.0000	2
Spore DNA protection	0.0000	2
Protein secretion system, Type V	0.0000	2
Coenzyme B	-0.0001	2
Reverse electron transport	-0.0001	2
Actinorhodin biosynthetic cluster	-0.0003	2
Proteasome related clusters	-0.0004	2
Riboflavin, FMN, FAD	-0.0013	2
Folate and pterines	-0.0016	2
Proteolytic pathway	-0.0017	2
Tricarboxylate transporter	-0.0018	2
pH adaptation potassium efflux	-0.0023	2
Signal transduction in Eukaryotes	-0.0025	2
DNA pol III alpha and a number of apparently unrelated functions	-0.0032	2
Cation transporters	-0.0033	2
Cell wall of Mycobacteria	-0.0037	2
Inorganic sulfur assimilation	-0.0043	2
Protein degradation	-0.0045	2
Degradation of Polyphenols (?)	-0.0051	2
Uni- Sym- and Antiporters	-0.0056	2
One-carbon Metabolism	-0.0058	2
Cold shock	-0.0069	2
Heat shock	-0.0083	2
Sodium Ion-Coupled Energetics	-0.0085	2
Hydrocarbons	-0.0093	2
Phages, Prophages, Transposable elements	-0.0108	2
Heterocyst formation	-0.0113	2
Sugar Phosphotransferase Systems, PTS	-0.0115	2
Cell Division	-0.0116	2
Anaerobic degradation of aromatic compounds	-0.0132	2
Plant-Prokaryote DOE project	-0.0135	2
Gram-Positive cell wall components	-0.0151	2
DNA repair	-0.0161	2
Cytochrome biogenesis	-0.0183	2
Catabolism of an unknown compound	-0.0216	2
TRAP transporters	-0.0217	2
Dessication stress	-0.0227	2
Coenzyme M	-0.0255	2
CO_2 fixation	-0.0282	2

Invasion and intracellular resistance	-0.0292	2
Plant Hormones	-0.0310	2
Quorum sensing and biofilm formation	-0.0312	2
RNA processing and modification	-0.0327	2
Protein processing and modification	-0.0332	2
DNA replication	-0.0387	2
Protein translocation across cytoplasmic membrane	-0.0437	2
Recombination related cluster	-0.0466	2
Pathogenicity islands	-0.0496	2
Protein folding	-0.0503	2
Electron transport and photophosphorylation	-0.0548	2
Light-harvesting complexes	-0.0563	2
Electron donating reactions	-0.0575	2
Phage family-specific subsystems	-0.0611	2
proteosome related	-0.0677	2
Transcription	-0.0992	2
Phages, Prophages	-0.1159	2
Biologically active compounds in metazoan cell defence and differentiation	-0.1192	2
Protein biosynthesis	-0.1749	2

Table III.S2 Metagenomic sample information, summary of RGI output for hits above mapping threshold (MAPQ>10 and minimum of 50 gene percent coverage) and proportion of sample reads mapped to CARD (ARG reads) that mapped back to MAGs. Download complete table with taxonomic assignment and EPSPS predicted sequence here.

Sample id	Sample description	Total reads (R1+R2)	Total reads mapped to CARD (average MAPQ>10, gene coverage>=50%)	ARO terms of ARGs found in sample (MAPQ>10, gene coverage>=50%)	Proportion of ARG reads mapped back to MAGs
C1_08_17	pond C1 in day 1	3.79E+07	0	NA	NA
C1_08_23	pond C1 in day 7	3.28E+07	0	NA	NA
C1_08_31	pond C1 in day 15	3.05E+07	0	NA	NA
C1_09_15	pond C1 in day 30	3.16E+07	0	NA	NA
C1_09_20	pond C1 in day 35	2.92E+07	0	NA	NA
C1_09_23	pond C1 in day 38	2.35E+07	0	NA	NA
C1_09_26	pond C1 in day 41	2.68E+07	0	NA	NA
C1_09_28	pond C1 in day 43	2.99E+07	0	NA	NA
C1_09_30	pond C1 in day 45	2.85E+07	0	NA	NA
C1_10_04	pond C1 in day 49	3.32E+07	451	MexF, OXA-12, CEPH-A3, cphA4, Rm3 beta-lactamase, MCR-7.1	0.753
C1_10_12	pond C1 in day 57	2.99E+07	28503	MexB, MexE, MexF, OprN, mexI, THIN-B beta-lactamase, ceoB, mexW, TriC, OpmH, mexK, Rm3 beta-lactamase, Pseudomonas aeruginosa CpxR, MuxB	0.499
C4_08_17	pond C4 in day 1	3.57E+07	0	NA	NA
C4_08_23	pond C4 in day 7	3.26E+07	0	NA	NA
C4_08_31	pond C4 in day 15	2.45E+07	0	NA	NA
C4_09_15	pond C4 in day 30	3.16E+07	0	NA	NA
C4_09_20	pond C4 in day 35	3.46E+07	0	NA	NA
C4_09_23	pond C4 in day 38	3.11E+07	54	APH(3")-Ib, APH(6)-Id	0.821
C4_09_26	pond C4 in day 41	2.91E+07	0	NA	NA

C4_09_28	pond C4 in day 43	2.98E+07	20	APH(3")-Ib, APH(6)-Id	0.913
C4_09_30	pond C4 in day 45	2.93E+07	6	APH(6)-Id	0.838
C4_10_04	pond C4 in day 49	3.89E+07	2456	MexB, MexE, MexF, OXA-12, APH(6)-Id, ceoB, mexW, cphA2, cphA4, CepS beta-lactamase, OpmH, mexK, Rm3 beta-lactamase, Pseudomonas aeruginosa CpxR, MuxB, OXA-724	0.617
C4_10_12	pond C4 in day 57	3.27E+07	1579	MexF, THIN-B beta-lactamase, CAU-1 beta-lactamase, OXA-12, ceoB, CepS beta-lactamase, Rm3 beta-lactamase, MuxB	0.679
C8_08_17	pond C8 in day 1	3.78E+07	0	NA	NA
C8_08_23	pond C8 in day 7	2.36E+07	84	OXA-12, MCR-7.1, AQU-2, OXA-724	0.622
C8_08_31	pond C8 in day 15	2.92E+07	13556	MexE, MexF, OprN, mexI, APH(3")-Ib, APH(6)-Id, ceoB, mexW, TriC, OpmH, mexK, Pseudomonas aeruginosa CpxR, MuxB	0.939
C8_09_15	pond C8 in day 30	3.13E+07	0	NA	NA
C8_09_20	pond C8 in day 35	3.82E+07	1719	MexE, MexF, mexW, mexK, Pseudomonas aeruginosa CpxR, MuxB	0.815
C8_09_23	pond C8 in day 38	2.81E+07	417	MexF, APH(3")-Ib, APH(6)-Id	0.644
C8_09_26	pond C8 in day 41	3.25E+07	829	MexF, CAU-1 beta-lactamase, APH(3")-Ib, APH(6)-Id, ceoB	0.732
C8_09_28	pond C8 in day 43	2.98E+07	334	APH(3")-Ib, APH(6)-Id, ceoB	0.759
C8_09_30	pond C8 in day 45	3.40E+07	32	APH(3")-Ib, APH(6)-Id	0.846
C8_10_04	pond C8 in day 49	3.83E+07	2473	MexB, MexF, CAU-1 beta-lactamase, ceoB, MuxB, OXA-724	0.473
C8_10_12	pond C8 in day 57	2.92E+07	1391	MexF, APH(3")-Ib, ceoB, MuxB, OXA-724	0.599

D1_08_17	pond D1 in day 1	3.55E+07	0	NA	NA
D1_08_23	pond D1 in day 7	3.14E+07	0	NA	NA
D1_08_31	pond D1 in day 15	3.72E+07	0	NA	NA
D1_09_15	pond D1 in day 30	3.42E+07	0	NA	NA
D1_09_20	pond D1 in day 35	3.29E+07	0	NA	NA
D1_09_23	pond D1 in day 38	7.34E+07	0	NA	NA
D1_09_26	pond D1 in day 41	7.25E+07	0	NA	NA
D1_09_28	pond D1 in day 43	2.62E+07	0	NA	NA
D1_09_30	pond D1 in day 45	6.53E+07	0	NA	NA
D1_10_04	pond D1 in day 49	3.10E+07	796	THIN-B beta-lactamase, OXA-12, LRA-18, cphA4, Rm3 beta-lactamase, MCR-7.1	0.877
D1_10_12	pond D1 in day 57	3.31E+07	3024	MexB, MexF, APH(3'')-Ib, ceoB, mexW, Rm3 beta-lactamase, MuxB	0.450
D4_08_17	pond D4 in day 1	3.16E+07	0	NA	NA
D4_08_23	pond D4 in day 7	3.26E+07	0	NA	NA
D4_08_31	pond D4 in day 15	4.11E+07	0	NA	NA
D4_09_15	pond D4 in day 30	1.91E+07	0	NA	NA
D4_09_20	pond D4 in day 35	3.15E+07	0	NA	NA
D4_09_23	pond D4 in day 38	2.41E+07	0	NA	NA
D4_09_26	pond D4 in day 41	3.10E+07	0	NA	NA
D4_09_28	pond D4 in day 43	2.68E+07	0	NA	NA
D4_09_30	pond D4 in day 45	2.49E+07	0	NA	NA
D4_10_04	pond D4 in day 49	2.59E+07	796	MexF, THIN-B beta-lactamase, OXA-12, LRA-18, mexW, mexK, Rm3 beta-lactamase, OXA-724	0.785
D4_10_12	pond D4 in day 57	2.87E+07	1541	MexE, MexF, ceoB, mexW, Rm3 beta-lactamase, MuxB	0.530
D8_08_17	pond D8 in day 1	3.10E+07	0	NA	NA
D8_08_23	pond D8 in day 7	3.37E+07	108	OXA-12, cphA4, MCR-7.1, OXA-724	0.606

D8_08_31	pond D8 in day 15	4.15E+07	2304	MexF, APH(3")-Ib, APH(6)-Id, ceoB	0.456
D8_09_15	pond D8 in day 30	3.57E+07	0	NA	NA
D8_09_20	pond D8 in day 35	3.62E+07	0	NA	NA
D8_09_23	pond D8 in day 38	2.72E+07	11	APH(3")-Ib	0.756
D8_09_26	pond D8 in day 41	2.94E+07	6	APH(6)-Id	0.849
D8_09_28	pond D8 in day 43	2.89E+07	0	NA	NA
D8_09_30	pond D8 in day 45	2.89E+07	0	NA	NA
D8_10_04	pond D8 in day 49	3.12E+07	979	MexF, CGB-1 beta-lactamase, THIN-B beta-lactamase, OXA-12, ceoB, Rm3 beta-lactamase, OXA-724	0.754
D8_10_12	pond D8 in day 57	2.91E+07	970	MexF, CGB-1 beta-lactamase, APH(3")-Ib, APH(6)-Id, ceoB, mexK	0.674
E1_08_17	pond E1 in day 1	3.19E+07	0	NA	NA
E1_08_23	pond E1 in day 7	2.65E+07	0	NA	NA
E1_08_31	pond E1 in day 15	3.08E+07	0	NA	NA
E1_09_15	pond E1 in day 30	2.87E+07	0	NA	NA
E1_09_20	pond E1 in day 35	3.32E+07	0	NA	NA
E1_09_23	pond E1 in day 38	2.99E+07	0	NA	NA
E1_09_26	pond E1 in day 41	2.87E+07	0	NA	NA
E1_09_28	pond E1 in day 43	2.87E+07	0	NA	NA
E1_09_30	pond E1 in day 45	2.95E+07	0	NA	NA
E1_10_04	pond E1 in day 49	3.13E+07	0	NA	NA
E1_10_12	pond E1 in day 57	2.93E+07	0	NA	NA
H1_08_17	pond H1 in day 1	3.46E+07	0	NA	NA
H1_08_23	pond H1 in day 7	3.45E+07	62	ceoB	0.896
H1_08_31	pond H1 in day 15	2.96E+07	0	NA	NA
H1_09_15	pond H1 in day 30	3.64E+07	0	NA	NA
H1_09_20	pond H1 in day 35	3.37E+07	0	NA	NA

H1_09_23	pond H1 in day 38	3.07E+07	0	NA	NA
H1_09_26	pond H1 in day 41	2.74E+07	0	NA	NA
H1_09_28	pond H1 in day 43	2.96E+07	0	NA	NA
H1_09_30	pond H1 in day 45	2.93E+07	0	NA	NA
H1_10_04	pond H1 in day 49	3.38E+07	0	NA	NA
H1_10_12	pond H1 in day 57	2.89E+07	0	NA	NA

Table III.S3 MAG information, predicted EPSPS amino acid sequence, summary of ARGs and plasmids. For each predicted EPSPS sequence, the putative classification regarding glyphosate resistance is shown. The number of potential plasmid contigs and how many of these had ARGs annotated is also shown. Number of ARGs annotated to MAG contigs (total RGI strict hits) are provided in the last column.

MAG id	NCBI BioSample accession	Genome size (bp)	Percent completion	Percent redundancy	EPSPS putative classification	Number of plasmidial contigs (70% probability)	Plasmidial contigs with ARGs (RGI strict hits)	Total ARGs in MAG (RGI strict hits)
C1_MAG_00001	SAMN21888747	2865274	100.0	0.0	resistant	11	0	0
C1_MAG_00004	SAMN21888748	2975196	99.3	0.7	resistant	56	0	0
C1_MAG_00006	SAMN21888749	4357056	99.3	0.7	sensitive	5	0	2
C1_MAG_00008	SAMN21888750	6550573	100.0	1.4	resistant	7	0	0
C1_MAG_00009	SAMN21888751	3904780	98.6	0.0	sensitive	0	0	1
C1_MAG_00011	SAMN21888752	6398453	98.6	0.7	unclassified	4	0	1
C1_MAG_00012	SAMN21888753	2031134	97.8	0.0	sensitive	0	0	0
C1_MAG_00018	SAMN21888754	5360263	97.8	0.7	sensitive	9	0	0
C1_MAG_00023	SAMN21888755	3996478	98.6	1.4	sensitive	8	0	1
C1_MAG_00025	SAMN21888756	2664860	97.8	1.4	unclassified	1	0	0
C1_MAG_00026	SAMN21888757	7721187	98.6	2.2	sensitive	30	0	0
C1_MAG_00029	SAMN21888758	1658808	96.4	0.7	sensitive	0	0	0
C1_MAG_00031	SAMN21888759	2702274	95.7	0.0	unclassified	9	0	0
C1_MAG_00033	SAMN21888760	5331647	97.1	2.2	resistant	6	0	2
C1_MAG_00036	SAMN21888761	3872062	94.2	0.7	sensitive	11	0	0
C1_MAG_00038	SAMN21888762	3462678	99.3	5.8	sensitive	4	1	2
C1_MAG_00039	SAMN21888763	6501613	93.5	1.4	resistant	57	0	2
C1_MAG_00042	SAMN21888764	4825613	93.5	4.3	sensitive	12	0	1
C1_MAG_00043	SAMN21888765	4064063	89.9	0.7	sensitive	113	0	2
C1_MAG_00044	SAMN21888766	3833755	88.5	0.7	sensitive	25	0	1

C1_MAG_00045	SAMN21888767	3486936	88.5	1.4	unclassified	56	0	0
C1_MAG_00046	SAMN21888768	3778426	87.1	0.0	sensitive	14	0	1
C1_MAG_00047	SAMN21888769	2520768	93.5	6.5	sensitive	90	0	0
C1_MAG_00048	SAMN21888770	2726274	93.5	6.5	sensitive	22	0	1
C1_MAG_00049	SAMN21888771	2709801	87.1	0.7	unclassified	0	0	0
C1_MAG_00050	SAMN21888772	5487486	85.6	0.0	resistant	53	1	1
C1_MAG_00051	SAMN21888773	3842240	87.8	2.2	sensitive	69	0	0
C1_MAG_00052	SAMN21888774	4613471	86.3	0.7	sensitive	1	0	5
C1_MAG_00054	SAMN21888775	3627059	84.9	0.7	sensitive	1	0	2
C1_MAG_00057	SAMN21888776	4080027	85.6	2.9	resistant	38	0	0
C1_MAG_00058	SAMN21888777	1550282	84.2	1.4	unclassified	3	0	0
C1_MAG_00060	SAMN21888778	4515140	87.8	5.8	sensitive	59	1	2
C1_MAG_00061	SAMN21888779	5804248	84.9	2.9	unclassified	128	0	2
C1_MAG_00064	SAMN21888781	2660907	80.6	1.4	unclassified	0	0	0
C1_MAG_00065	SAMN21888782	2673836	85.6	7.9	unclassified	19	0	0
C1_MAG_00066	SAMN21888783	2028447	75.5	0.0	sensitive	19	0	1
C1_MAG_00067	SAMN21888784	3571564	77.7	2.2	unclassified	275	0	0
C1_MAG_00068	SAMN21888785	4731322	77.0	2.2	resistant	21	0	1
C1_MAG_00069	SAMN21888786	2038476	76.3	2.2	sensitive	122	0	0
C1_MAG_00071	SAMN21888787	2808021	75.5	2.2	sensitive	279	0	0
C1_MAG_00072	SAMN21888788	3099470	76.3	2.9	unclassified	71	0	1
C1_MAG_00073	SAMN21888789	3770674	79.1	5.8	resistant	55	0	0
C1_MAG_00074	SAMN21888790	1132460	74.8	3.6	unclassified	7	0	0
C1_MAG_00075	SAMN21888791	6030439	72.7	2.2	sensitive	74	0	1
C1_MAG_00076	SAMN21888792	5707845	77.0	6.5	resistant	88	0	1
C1_MAG_00077	SAMN21888793	2546518	74.1	4.3	resistant	8	0	0
C1_MAG_00078	SAMN21888794	1867033	71.2	2.2	resistant	13	0	0
C4_MAG_00008	SAMN21888795	1861907	97.8	0.0	unclassified	4	0	1

C4_MAG_00010	SAMN21888796	3314809	97.1	0.0	sensitive	25	0	0
C4_MAG_00011	SAMN21888797	3340832	97.1	0.0	sensitive	1	0	0
C4_MAG_00013	SAMN21888798	3650414	98.6	1.4	unclassified	10	0	0
C4_MAG_00014	SAMN21888799	2370615	97.1	0.0	resistant	0	0	1
C4_MAG_00015	SAMN21888800	4080143	98.6	2.2	sensitive	10	0	0
C4_MAG_00016	SAMN21888801	1795656	96.4	0.0	unclassified	3	0	0
C4_MAG_00017	SAMN21888802	1525279	96.4	0.0	unclassified	0	0	0
C4_MAG_00019	SAMN21888803	4399728	96.4	0.7	resistant	14	0	1
C4_MAG_00020	SAMN21888804	2865212	97.8	2.2	unclassified	26	0	0
C4_MAG_00023	SAMN21888805	3451066	98.6	3.6	sensitive	9	0	0
C4_MAG_00025	SAMN21888806	3239000	96.4	2.2	sensitive	26	0	2
C4_MAG_00027	SAMN21888807	2786827	95.7	2.2	sensitive	10	0	0
C4_MAG_00028	SAMN21888808	3158231	96.4	2.9	resistant	13	0	0
C4_MAG_00029	SAMN21888809	1673851	93.5	0.0	sensitive	25	0	1
C4_MAG_00031	SAMN21888810	4053606	95.0	2.2	unclassified	0	0	0
C4_MAG_00032	SAMN21888811	6591471	96.4	3.6	unclassified	1	0	0
C4_MAG_00033	SAMN21888812	3486739	98.6	5.8	sensitive	1	0	0
C4_MAG_00035	SAMN21888813	3187320	93.5	0.7	resistant	0	0	0
C4_MAG_00036	SAMN21888814	3453136	96.4	4.3	resistant	22	0	2
C4_MAG_00039	SAMN21888815	3573516	94.2	2.2	sensitive	201	1	1
C4_MAG_00041	SAMN21888816	3906659	92.1	0.7	sensitive	16	0	0
C4_MAG_00042	SAMN21888817	2474022	92.8	1.4	sensitive	34	0	1
C4_MAG_00043	SAMN21888818	4842440	93.5	2.2	unclassified	125	1	1
C4_MAG_00045	SAMN21888819	2004405	91.4	0.0	sensitive	4	0	0
C4_MAG_00046	SAMN21888820	3323307	92.8	1.4	resistant	12	0	0
C4_MAG_00047	SAMN21888821	2979333	95.7	5.0	resistant	31	0	0
C4_MAG_00048	SAMN21888822	1703749	91.4	0.7	sensitive	23	0	0
C4_MAG_00049	SAMN21888823	3393067	92.8	2.9	resistant	66	0	0

C4_MAG_00051	SAMN21888824	3389472	96.4	7.2	unclassified	38	0	0
C4_MAG_00053	SAMN21888825	2539435	90.6	2.2	unclassified	36	0	0
C4_MAG_00054	SAMN21888826	1272719	90.6	2.2	unclassified	1	0	0
C4_MAG_00055	SAMN21888827	2632590	90.6	2.2	resistant	14	0	0
C4_MAG_00056	SAMN21888828	3243781	89.9	2.9	resistant	15	0	0
C4_MAG_00057	SAMN21888829	2352710	88.5	2.2	sensitive	10	0	1
C4_MAG_00058	SAMN21888830	1675113	85.6	0.0	sensitive	1	0	2
C4_MAG_00059	SAMN21888831	3189848	85.6	0.7	sensitive	7	0	1
C4_MAG_00060	SAMN21888832	6072898	86.3	1.4	resistant	40	1	2
C4_MAG_00061	SAMN21888833	3704697	88.5	4.3	unclassified	40	0	3
C4_MAG_00062	SAMN21888834	1645787	85.6	1.4	sensitive	3	0	0
C4_MAG_00063	SAMN21888835	4629015	84.9	1.4	sensitive	6	0	2
C4_MAG_00064	SAMN21888836	3981732	87.8	4.3	resistant	82	0	0
C4_MAG_00066	SAMN21888837	5516459	84.9	1.4	resistant	50	1	2
C4_MAG_00067	SAMN21888838	3726655	89.9	6.5	resistant	28	0	0
C4_MAG_00068	SAMN21888839	2838606	85.6	2.9	sensitive	25	0	0
C4_MAG_00069	SAMN21888840	2055836	84.9	2.2	sensitive	14	0	1
C4_MAG_00070	SAMN21888841	2309648	88.5	5.8	sensitive	15	0	0
C4_MAG_00072	SAMN21888842	3454674	85.6	4.3	sensitive	6	0	0
C4_MAG_00073	SAMN21888843	5579033	82.0	0.7	unclassified	4	0	3
C4_MAG_00077	SAMN21888844	3636379	82.0	2.2	resistant	28	0	0
C4_MAG_00079	SAMN21888845	1579574	80.6	1.4	unclassified	1	0	0
C4_MAG_00080	SAMN21888846	796455	77.7	0.0	unclassified	2	0	0
C4_MAG_00081	SAMN21888847	2277307	77.7	0.0	sensitive	1	0	0
C4_MAG_00083	SAMN21888848	3835135	79.9	2.9	resistant	140	0	0
C4_MAG_00085	SAMN21888849	3075021	79.1	4.3	sensitive	58	0	2
C4_MAG_00087	SAMN21888850	4107736	77.7	3.6	resistant	57	0	0
C4_MAG_00088	SAMN21888851	4061822	75.5	4.3	unclassified	18	0	1

C4_MAG_00089	SAMN21888852	5663909	72.7	2.2	sensitive	0	0	5
C4_MAG_00090	SAMN21888853	4321013	75.5	5.8	resistant	356	0	1
C4_MAG_00091	SAMN21888854	1539032	72.7	3.6	unclassified	65	0	0
C4_MAG_00092	SAMN21888855	2702925	71.9	3.6	unclassified	30	0	2
C4_MAG_00096	SAMN21888856	1336791	71.9	7.2	unclassified	12	0	0
C8_MAG_00001	SAMN21888857	4055180	100.0	0.0	resistant	10	1	2
C8_MAG_00002	SAMN21888858	3133556	100.0	0.0	sensitive	2	0	0
C8_MAG_00003	SAMN21888859	3349211	100.0	0.0	sensitive	0	0	0
C8_MAG_00006	SAMN21888860	3594080	99.3	0.0	resistant	13	0	0
C8_MAG_00007	SAMN21888861	4330577	99.3	0.7	sensitive	4	0	2
C8_MAG_00008	SAMN21888862	4525610	99.3	1.4	sensitive	1	0	0
C8_MAG_00010	SAMN21888863	2964701	100.0	2.9	resistant	74	0	0
C8_MAG_00011	SAMN21888864	4403135	99.3	2.9	sensitive	0	0	2
C8_MAG_00012	SAMN21888865	4009353	99.3	2.9	resistant	58	0	1
C8_MAG_00013	SAMN21888866	2355252	97.1	1.4	sensitive	0	0	2
C8_MAG_00014	SAMN21888867	7122131	97.1	1.4	sensitive	17	0	0
C8_MAG_00015	SAMN21888868	1806409	96.4	0.7	sensitive	0	0	0
C8_MAG_00016	SAMN21888869	4661110	98.6	2.9	resistant	186	2	2
C8_MAG_00017	SAMN21888870	4322018	96.4	2.2	sensitive	43	1	4
C8_MAG_00018	SAMN21888871	4632588	99.3	5.0	unclassified	12	0	0
C8_MAG_00019	SAMN21888872	5881872	96.4	2.9	resistant	45	3	3
C8_MAG_00020	SAMN21888873	4579689	96.4	3.6	unclassified	4	0	0
C8_MAG_00021	SAMN21888874	3704676	94.2	1.4	sensitive	19	0	2
C8_MAG_00023	SAMN21888875	3473837	93.5	1.4	sensitive	4	0	2
C8_MAG_00024	SAMN21888876	3409116	99.3	7.2	unclassified	6	0	1
C8_MAG_00028	SAMN21888877	4006252	94.2	5.8	unclassified	137	1	1
C8_MAG_00029	SAMN21888878	4038507	94.2	5.8	resistant	20	0	1
C8_MAG_00031	SAMN21888879	4695849	92.1	4.3	resistant	78	0	1

C8_MAG_00032	SAMN21888880	3976838	91.4	3.6	unclassified	72	0	2
C8_MAG_00033	SAMN21888881	3065071	84.9	0.0	sensitive	112	0	1
C8_MAG_00034	SAMN21888882	4744953	92.1	7.2	resistant	18	0	0
C8_MAG_00036	SAMN21888883	3360869	85.6	1.4	sensitive	1	0	0
C8_MAG_00037	SAMN21888884	5019011	85.6	1.4	resistant	45	0	0
C8_MAG_00038	SAMN21888885	6557524	86.3	2.9	resistant	1	0	4
C8_MAG_00039	SAMN21888886	3125813	84.2	0.7	unclassified	4	0	1
C8_MAG_00040	SAMN21888887	2353520	83.5	1.4	unclassified	87	0	0
C8_MAG_00041	SAMN21888888	5515789	84.2	2.9	resistant	116	1	1
C8_MAG_00042	SAMN21888889	3150239	82.0	0.7	resistant	2	0	1
C8_MAG_00044	SAMN21888890	3952081	83.5	2.9	unclassified	314	0	1
C8_MAG_00046	SAMN21888891	3975225	84.9	7.2	resistant	58	0	0
C8_MAG_00048	SAMN21888892	5194788	79.9	4.3	resistant	229	0	2
C8_MAG_00050	SAMN21888893	2849320	79.1	5.0	resistant	25	0	1
C8_MAG_00051	SAMN21888894	2634691	75.5	1.4	unclassified	9	0	0
C8_MAG_00052	SAMN21888895	3566969	80.6	6.5	resistant	520	0	0
C8_MAG_00054	SAMN21888896	1289552	73.4	0.7	unclassified	0	0	0
C8_MAG_00055	SAMN21888897	1517273	72.7	0.0	unclassified	45	0	0
C8_MAG_00057	SAMN21888898	4818823	74.1	5.8	sensitive	20	0	2
C8_MAG_00058	SAMN21888899	2948086	72.7	5.0	resistant	3	0	2
C8_MAG_00060	SAMN21888900	1425545	70.5	2.9	sensitive	2	0	0
D1_MAG_00001	SAMN21888901	3604720	100.0	0.7	sensitive	0	0	0
D1_MAG_00002	SAMN21888902	2896378	99.3	0.0	sensitive	5	0	2
D1_MAG_00006	SAMN21888903	4270577	98.6	0.0	unclassified	0	0	1
D1_MAG_00008	SAMN21888904	2394771	97.8	0.0	unclassified	4	0	0
D1_MAG_00009	SAMN21888905	4171199	100.0	2.2	resistant	28	0	0
D1_MAG_00011	SAMN21888906	3102269	99.3	2.2	sensitive	3	0	1
D1_MAG_00012	SAMN21888907	4966842	99.3	2.2	resistant	27	0	0

D1_MAG_00013	SAMN21888908	3331068	97.1	0.0	sensitive	8	0	0
D1_MAG_00014	SAMN21888909	4833864	97.8	1.4	sensitive	0	0	4
D1_MAG_00017	SAMN21888910	4106956	97.8	1.4	sensitive	9	0	3
D1_MAG_00018	SAMN21888911	3777048	98.6	2.2	unclassified	21	0	1
D1_MAG_00019	SAMN21888912	5046364	97.1	0.7	resistant	156	0	1
D1_MAG_00021	SAMN21888913	3774278	95.7	0.0	resistant	0	0	0
D1_MAG_00028	SAMN21888914	2664058	95.7	1.4	unclassified	21	0	0
D1_MAG_00033	SAMN21888915	3902075	93.5	0.0	sensitive	33	0	1
D1_MAG_00034	SAMN21888916	3582211	94.2	0.7	resistant	12	0	0
D1_MAG_00036	SAMN21888917	4398125	95.7	3.6	resistant	136	0	0
D1_MAG_00037	SAMN21888918	4255971	97.1	5.8	sensitive	23	0	0
D1_MAG_00039	SAMN21888919	1941960	92.8	1.4	unclassified	8	0	1
D1_MAG_00040	SAMN21888920	5357856	95.0	4.3	sensitive	223	1	2
D1_MAG_00042	SAMN21888921	5080471	97.8	7.2	sensitive	21	0	1
D1_MAG_00043	SAMN21888922	2868958	96.4	5.8	resistant	46	0	1
D1_MAG_00046	SAMN21888923	3175710	98.6	8.6	sensitive	37	0	2
D1_MAG_00047	SAMN21888924	3166428	91.4	1.4	resistant	22	0	0
D1_MAG_00048	SAMN21888925	3075948	91.4	2.2	unclassified	4	0	0
D1_MAG_00049	SAMN21888926	5137330	96.4	7.2	resistant	62	0	2
D1_MAG_00050	SAMN21888927	5563704	92.1	3.6	resistant	437	0	3
D1_MAG_00051	SAMN21888928	3755658	93.5	5.0	resistant	24	0	1
D1_MAG_00052	SAMN21888929	5814712	92.8	4.3	resistant	14	0	0
D1_MAG_00053	SAMN21888930	5863037	89.2	1.4	sensitive	79	0	3
D1_MAG_00054	SAMN21888931	2824473	91.4	4.3	resistant	92	0	0
D1_MAG_00055	SAMN21888932	4962122	88.5	2.2	unclassified	349	0	0
D1_MAG_00056	SAMN21888933	4024280	91.4	5.0	resistant	37	0	2
D1_MAG_00057	SAMN21888934	3691626	89.9	3.6	unclassified	27	0	1
D1_MAG_00059	SAMN21888935	2857864	87.1	2.9	sensitive	144	0	0

D1_MAG_00060	SAMN21888936	3320411	87.8	3.6	unclassified	30	0	1
D1_MAG_00064	SAMN21888937	2800134	83.5	1.4	resistant	16	0	1
D1_MAG_00065	SAMN21888938	3003553	84.2	2.9	resistant	114	0	0
D1_MAG_00066	SAMN21888939	3798204	82.7	1.4	resistant	277	0	1
D1_MAG_00067	SAMN21888940	5232909	84.2	4.3	resistant	218	1	1
D1_MAG_00068	SAMN21888941	4351049	86.3	7.2	unclassified	15	0	0
D1_MAG_00070	SAMN21888942	4338921	78.4	0.0	sensitive	3	0	2
D1_MAG_00071	SAMN21888943	6597014	86.3	8.6	sensitive	243	0	2
D1_MAG_00073	SAMN21888944	2890219	82.0	4.3	resistant	42	0	0
D1_MAG_00075	SAMN21888945	1732735	77.0	0.0	unclassified	51	0	0
D1_MAG_00076	SAMN21888946	732660	77.0	0.0	unclassified	1	0	0
D1_MAG_00080	SAMN21888948	4044923	81.3	7.9	sensitive	5	0	1
D1_MAG_00081	SAMN21888949	6661811	80.6	7.9	sensitive	3	0	3
D1_MAG_00082	SAMN21888950	3352437	71.9	1.4	unclassified	264	0	0
D1_MAG_00083	SAMN21888951	4278001	73.4	5.0	sensitive	9	0	0
D1_MAG_00084	SAMN21888952	2523862	71.2	3.6	sensitive	20	0	2
D4_MAG_00001	SAMN21888953	2593686	99.3	0.0	sensitive	2	0	0
D4_MAG_00002	SAMN21888954	4018767	98.6	0.0	sensitive	3	0	0
D4_MAG_00003	SAMN21888955	4217982	99.3	0.7	sensitive	0	0	2
D4_MAG_00005	SAMN21888956	2723044	99.3	0.7	sensitive	0	0	0
D4_MAG_00008	SAMN21888957	1804560	100.0	2.2	unclassified	8	0	1
D4_MAG_00009	SAMN21888958	6425337	98.6	0.7	sensitive	2	0	0
D4_MAG_00010	SAMN21888959	6387458	100.0	2.9	resistant	14	0	1
D4_MAG_00012	SAMN21888960	2505487	97.8	0.7	sensitive	8	0	0
D4_MAG_00013	SAMN21899414	2442503	97.8	0.7	sensitive	1	0	1
D4_MAG_00014	SAMN21899415	2519991	97.1	0.0	sensitive	2	0	1
D4_MAG_00015	SAMN21899416	1446827	97.1	0.0	unclassified	0	0	0
D4_MAG_00016	SAMN21899417	2919762	97.8	0.7	sensitive	1	0	0

D4_MAG_00017	SAMN21899418	6500657	99.3	2.9	resistant	26	0	0
D4_MAG_00019	SAMN21899419	3431227	98.6	2.2	resistant	77	0	0
D4_MAG_00020	SAMN21899420	3391886	100.0	4.3	resistant	12	0	0
D4_MAG_00021	SAMN21899421	3167197	97.1	2.2	unclassified	10	0	0
D4_MAG_00022	SAMN21899422	2488393	95.7	0.7	sensitive	1	0	0
D4_MAG_00024	SAMN21899423	5729541	97.1	2.9	resistant	5	0	2
D4_MAG_00025	SAMN21899424	2124332	94.2	0.0	sensitive	2	0	0
D4_MAG_00026	SAMN21899425	2954225	98.6	5.0	sensitive	30	0	2
D4_MAG_00030	SAMN21899426	1923915	96.4	2.9	sensitive	30	0	2
D4_MAG_00031	SAMN21899427	2547609	93.5	0.0	unclassified	14	0	0
D4_MAG_00032	SAMN21899428	2991422	94.2	0.7	sensitive	0	0	0
D4_MAG_00033	SAMN21899429	3021296	94.2	0.7	unclassified	1	0	0
D4_MAG_00034	SAMN21899430	3465646	95.7	2.9	unclassified	2	0	0
D4_MAG_00035	SAMN21899431	3666610	95.7	2.9	sensitive	12	0	1
D4_MAG_00038	SAMN21899432	3732554	94.2	1.4	resistant	34	0	1
D4_MAG_00039	SAMN21899433	2962799	92.1	0.0	resistant	5	0	0
D4_MAG_00040	SAMN21899434	3491904	94.2	2.2	unclassified	27	0	1
D4_MAG_00042	SAMN21899435	1754175	92.1	0.7	sensitive	2	0	0
D4_MAG_00043	SAMN21899436	1416337	92.8	1.4	unclassified	26	0	1
D4_MAG_00044	SAMN21899437	2766400	91.4	0.0	sensitive	4	0	1
D4_MAG_00045	SAMN21899438	2567491	95.7	5.0	sensitive	3	0	0
D4_MAG_00047	SAMN21899439	3501432	91.4	1.4	resistant	20	0	1
D4_MAG_00048	SAMN21899440	2328478	91.4	1.4	resistant	6	0	0
D4_MAG_00049	SAMN21899441	1234801	88.5	0.0	sensitive	3	0	0
D4_MAG_00050	SAMN21899442	2610438	88.5	0.7	resistant	36	0	0
D4_MAG_00051	SAMN21899443	2055789	88.5	1.4	sensitive	6	0	2
D4_MAG_00052	SAMN21899444	4453797	92.1	5.8	unclassified	22	0	0
D4_MAG_00054	SAMN21899445	5143557	87.1	1.4	sensitive	4	0	4

D4_MAG_00056	SAMN21899446	3490039	92.1	7.9	sensitive	123	0	0
D4_MAG_00057	SAMN21899447	2791729	83.5	0.0	unclassified	18	0	0
D4_MAG_00060	SAMN21899448	4856506	91.4	7.9	unclassified	58	0	5
D4_MAG_00061	SAMN21899449	1784527	84.2	1.4	unclassified	11	0	0
D4_MAG_00063	SAMN21899450	2709729	86.3	5.0	sensitive	29	0	0
D4_MAG_00065	SAMN21899451	2781385	84.2	4.3	sensitive	62	1	3
D4_MAG_00068	SAMN21899452	1497840	77.7	0.0	unclassified	41	0	0
D4_MAG_00069	SAMN21899453	772463	77.0	0.0	unclassified	12	0	0
D4_MAG_00072	SAMN21899454	2823730	75.5	0.7	sensitive	11	0	1
D4_MAG_00074	SAMN21899455	1964383	78.4	5.0	sensitive	54	0	0
D4_MAG_00075	SAMN21899456	4816134	74.8	2.2	resistant	24	0	2
D4_MAG_00076	SAMN21899457	3027117	74.8	2.9	resistant	24	0	2
D4_MAG_00077	SAMN21899458	1638219	76.3	4.3	unclassified	17	0	1
D4_MAG_00078	SAMN21899459	2941426	71.9	0.0	resistant	163	0	0
D4_MAG_00079	SAMN21899460	1802556	72.7	1.4	unclassified	7	0	0
D4_MAG_00080	SAMN21899461	1102693	72.7	1.4	unclassified	33	0	0
D4_MAG_00082	SAMN21899462	2294988	71.2	3.6	unclassified	16	0	1
D4_MAG_00083	SAMN21899463	3367231	71.2	5.0	resistant	61	0	0
D4_MAG_00084	SAMN21899464	1890213	73.4	7.2	unclassified	13	0	1
D8_MAG_00001	SAMN21899465	3382297	100.0	0.0	sensitive	2	0	0
D8_MAG_00002	SAMN21899466	3491239	99.3	0.0	sensitive	3	0	0
D8_MAG_00007	SAMN21899467	4598917	97.8	0.0	sensitive	5	0	4
D8_MAG_00011	SAMN21899468	3767514	97.8	0.7	unclassified	0	0	0
D8_MAG_00014	SAMN21899469	2869542	97.8	1.4	sensitive	3	0	0
D8_MAG_00015	SAMN21899470	3092798	99.3	2.9	resistant	33	0	2
D8_MAG_00016	SAMN21899471	6981573	97.8	1.4	resistant	27	0	3
D8_MAG_00017	SAMN21899472	3509968	98.6	2.2	sensitive	9	0	3
D8_MAG_00019	SAMN21899473	3990259	95.7	0.0	sensitive	2	0	0

D8_MAG_00020	SAMN21899474	2655977	95.7	0.7	unclassified	0	0	1
D8_MAG_00021	SAMN21899475	3057546	95.7	0.7	resistant	23	0	0
D8_MAG_00022	SAMN21899476	3160708	98.6	3.6	sensitive	3	0	2
D8_MAG_00024	SAMN21899477	3151776	95.0	2.2	resistant	8	0	2
D8_MAG_00026	SAMN21899478	2949222	95.7	3.6	resistant	46	0	0
D8_MAG_00028	SAMN21899479	3017821	94.2	2.2	sensitive	4	0	0
D8_MAG_00029	SAMN21899480	1894213	92.1	0.0	sensitive	8	0	1
D8_MAG_00031	SAMN21899481	3635657	95.0	4.3	unclassified	48	0	0
D8_MAG_00032	SAMN21899482	3288821	92.8	2.2	resistant	42	0	0
D8_MAG_00033	SAMN21899483	3130975	90.6	0.0	unclassified	0	0	1
D8_MAG_00034	SAMN21899484	6511187	92.1	1.4	resistant	84	1	3
D8_MAG_00035	SAMN21899485	1789251	92.1	2.2	resistant	2	0	0
D8_MAG_00036	SAMN21899486	3096660	91.4	1.4	resistant	44	0	0
D8_MAG_00037	SAMN21899487	5938020	93.5	4.3	sensitive	102	0	0
D8_MAG_00040	SAMN21899488	6199988	91.4	2.2	sensitive	64	0	1
D8_MAG_00041	SAMN21899489	3410854	89.2	2.2	unclassified	172	0	1
D8_MAG_00044	SAMN21899490	4438271	89.9	3.6	resistant	34	0	0
D8_MAG_00045	SAMN21899491	1733669	87.1	1.4	sensitive	19	0	1
D8_MAG_00046	SAMN21899492	2842418	86.3	0.7	sensitive	148	0	1
D8_MAG_00048	SAMN21899493	4649363	90.6	5.8	resistant	69	0	0
D8_MAG_00049	SAMN21899494	3429808	86.3	1.4	resistant	81	1	1
D8_MAG_00050	SAMN21899495	3834122	90.6	5.8	resistant	41	0	0
D8_MAG_00051	SAMN21899496	5928705	86.3	2.2	sensitive	2	0	3
D8_MAG_00053	SAMN21899497	2938286	87.8	4.3	sensitive	28	0	1
D8_MAG_00054	SAMN21899498	3424644	86.3	2.9	resistant	3	0	1
D8_MAG_00055	SAMN21899499	3275666	83.5	0.7	unclassified	267	0	0
D8_MAG_00056	SAMN21899500	1837119	85.6	3.6	unclassified	24	0	1
D8_MAG_00057	SAMN21899501	4474161	86.3	5.0	sensitive	6	0	1

D8_MAG_00058	SAMN21899502	4286318	85.6	4.3	resistant	33	0	1
D8_MAG_00059	SAMN21899503	4149274	82.0	0.7	unclassified	4	0	1
D8_MAG_00060	SAMN21899504	1597893	83.5	2.9	unclassified	8	0	1
D8_MAG_00061	SAMN21899505	3939285	82.0	1.4	resistant	79	0	1
D8_MAG_00063	SAMN21899506	3885329	80.6	1.4	resistant	10	0	0
D8_MAG_00064	SAMN21899507	3103980	79.9	0.7	unclassified	0	0	1
D8_MAG_00066	SAMN21899508	4207543	79.9	2.2	resistant	56	0	0
D8_MAG_00068	SAMN21899509	3344199	80.6	4.3	sensitive	0	0	1
D8_MAG_00069	SAMN21899510	5104509	83.5	7.2	resistant	61	0	1
D8_MAG_00071	SAMN21899511	4365470	77.0	1.4	resistant	150	0	1
D8_MAG_00072	SAMN21899512	3512356	76.3	0.7	sensitive	1	0	1
D8_MAG_00073	SAMN21899513	5043564	84.2	8.6	resistant	67	1	1
D8_MAG_00074	SAMN21899514	3770652	74.1	0.0	resistant	18	0	0
D8_MAG_00075	SAMN21899515	3242775	75.5	2.2	sensitive	1	0	0
D8_MAG_00076	SAMN21899516	4180664	73.4	0.0	resistant	12	0	1
D8_MAG_00078	SAMN21899517	3796550	74.8	2.9	resistant	146	0	0
D8_MAG_00079	SAMN21899518	5925387	72.7	1.4	resistant	158	0	1
D8_MAG_00080	SAMN21899519	5196463	73.4	2.2	unclassified	330	0	0
D8_MAG_00082	SAMN21899520	2636802	71.9	1.4	sensitive	45	0	1
D8_MAG_00083	SAMN21899521	3317900	71.2	1.4	resistant	48	0	0
D8_MAG_00085	SAMN21899522	3376023	77.7	9.4	resistant	8	1	2
E1_MAG_00001	SAMN21899523	3035998	99.3	0.7	unclassified	2	0	0
E1_MAG_00003	SAMN21899524	3282591	99.3	1.4	resistant	8	0	0
E1_MAG_00004	SAMN21899525	6277322	99.3	1.4	sensitive	0	0	0
E1_MAG_00006	SAMN21899526	2698210	99.3	2.2	resistant	1	0	0
E1_MAG_00007	SAMN21899527	3561788	97.1	0.0	sensitive	3	0	1
E1_MAG_00009	SAMN21899528	2423889	99.3	2.2	sensitive	3	0	1
E1_MAG_00012	SAMN21899529	3551629	97.8	1.4	sensitive	0	0	0

E1_MAG_00013	SAMN21899530	4326669	97.8	2.2	sensitive	5	0	0
E1_MAG_00014	SAMN21899531	8365988	98.6	2.9	sensitive	11	0	0
E1_MAG_00016	SAMN21899532	2492545	97.8	2.2	sensitive	0	0	0
E1_MAG_00017	SAMN21899533	3374941	95.7	0.0	sensitive	11	0	0
E1_MAG_00018	SAMN21899534	2188583	98.6	2.9	resistant	5	0	0
E1_MAG_00022	SAMN21899535	3481644	97.8	2.9	resistant	2	0	2
E1_MAG_00024	SAMN21899536	2718561	95.0	0.7	unclassified	0	0	0
E1_MAG_00026	SAMN21899537	2497358	96.4	2.9	sensitive	2	0	1
E1_MAG_00027	SAMN21899538	5390984	98.6	5.0	unclassified	2	0	1
E1_MAG_00028	SAMN21899539	2162404	95.0	1.4	unclassified	3	0	0
E1_MAG_00030	SAMN21899540	2859717	96.4	2.9	unclassified	116	0	0
E1_MAG_00031	SAMN21899541	2673417	95.0	1.4	unclassified	50	0	0
E1_MAG_00033	SAMN21899542	5014731	97.1	4.3	resistant	11	0	1
E1_MAG_00037	SAMN21899543	2270111	92.8	1.4	resistant	23	0	0
E1_MAG_00038	SAMN21899544	4133991	92.1	0.7	sensitive	36	0	0
E1_MAG_00040	SAMN21899545	1156315	91.4	1.4	unclassified	0	0	0
E1_MAG_00042	SAMN21899546	2705774	92.1	2.2	sensitive	44	1	3
E1_MAG_00043	SAMN21899547	1358893	91.4	2.9	sensitive	16	0	0
E1_MAG_00046	SAMN21899548	4892287	90.6	5.0	unclassified	165	0	0
E1_MAG_00048	SAMN21899549	2742593	93.5	7.9	unclassified	1	0	0
E1_MAG_00050	SAMN21899550	2213866	84.9	1.4	unclassified	11	0	0
E1_MAG_00051	SAMN21899551	2861209	86.3	2.9	sensitive	4	0	0
E1_MAG_00052	SAMN21899552	5389703	85.6	2.2	sensitive	236	0	1
E1_MAG_00053	SAMN21899553	2714334	84.9	1.4	sensitive	43	0	0
E1_MAG_00054	SAMN21899554	2214136	84.9	2.2	unclassified	2	0	1
E1_MAG_00055	SAMN21899555	2449389	84.2	2.2	sensitive	0	0	2
E1_MAG_00056	SAMN21899556	5010097	84.2	2.2	sensitive	25	0	0
E1_MAG_00057	SAMN21899557	1302914	82.0	0.0	sensitive	6	0	0

E1_MAG_00058	SAMN21899558	3218503	82.7	1.4	unclassified	2	0	0
E1_MAG_00060	SAMN21899559	2205660	82.7	2.9	unclassified	3	0	0
E1_MAG_00061	SAMN21899560	1948469	80.6	1.4	unclassified	5	0	0
E1_MAG_00062	SAMN21899561	2616823	83.5	5.0	sensitive	2	0	1
E1_MAG_00063	SAMN21899562	3309107	83.5	5.0	unclassified	39	0	1
E1_MAG_00064	SAMN21899563	3073038	81.3	4.3	sensitive	4	0	1
E1_MAG_00065	SAMN21899564	3869542	80.6	5.0	resistant	61	0	0
E1_MAG_00066	SAMN21899565	4566525	82.7	7.9	resistant	61	0	2
E1_MAG_00067	SAMN21899566	2664125	74.8	0.7	unclassified	43	0	1
E1_MAG_00068	SAMN21899567	1141766	78.4	5.0	sensitive	4	0	0
E1_MAG_00071	SAMN21899568	2954750	72.7	0.0	sensitive	10	0	0
E1_MAG_00072	SAMN21899569	4207763	75.5	2.9	sensitive	24	0	0
E1_MAG_00073	SAMN21899570	2344386	75.5	3.6	unclassified	33	0	2
E1_MAG_00074	SAMN21899571	4360381	78.4	7.2	sensitive	25	0	0
E1_MAG_00075	SAMN21899572	2756146	71.2	1.4	sensitive	55	0	0
E1_MAG_00076	SAMN21899573	2385757	71.9	2.2	resistant	56	0	0
E1_MAG_00077	SAMN21899574	2200563	71.9	2.9	resistant	73	0	0
E1_MAG_00078	SAMN21899575	3284141	75.5	8.6	sensitive	9	0	0
E1_MAG_00080	SAMN21899576	2761726	70.5	8.6	unclassified	107	0	0
H1_MAG_00001	SAMN21899577	2654371	99.3	0.0	resistant	13	0	0
H1_MAG_00003	SAMN21899578	3332312	100.0	0.7	resistant	4	0	0
H1_MAG_00006	SAMN21899579	4350866	99.3	1.4	resistant	13	0	2
H1_MAG_00009	SAMN21899580	2211795	97.1	0.0	sensitive	0	0	0
H1_MAG_00010	SAMN21899581	1465243	97.1	0.0	sensitive	3	0	0
H1_MAG_00011	SAMN21899582	3378966	98.6	1.4	sensitive	1	0	0
H1_MAG_00012	SAMN21899583	4547569	97.1	0.0	sensitive	3	0	1
H1_MAG_00013	SAMN21899584	3701345	97.8	0.7	resistant	4	0	2
H1_MAG_00014	SAMN21899585	3196577	98.6	2.2	resistant	15	0	0

H1_MAG_00017	SAMN21899586	4099764	97.1	0.7	sensitive	1	0	0
H1_MAG_00018	SAMN21899587	2633236	97.8	2.2	resistant	18	0	0
H1_MAG_00021	SAMN21899588	3670329	96.4	1.4	sensitive	11	0	1
H1_MAG_00022	SAMN21899589	2886961	96.4	1.4	sensitive	4	0	0
H1_MAG_00023	SAMN21899590	3788203	95.0	0.7	unclassified	66	0	0
H1_MAG_00024	SAMN21899591	5184536	97.1	2.9	resistant	72	0	1
H1_MAG_00025	SAMN21899592	4281394	97.8	3.6	unclassified	45	0	0
H1_MAG_00026	SAMN21899593	3545296	99.3	5.0	sensitive	4	0	2
H1_MAG_00027	SAMN21899594	6127118	97.8	4.3	sensitive	165	0	0
H1_MAG_00028	SAMN21899595	4428393	97.8	4.3	sensitive	11	0	0
H1_MAG_00030	SAMN21899596	4534679	94.2	1.4	resistant	18	0	0
H1_MAG_00031	SAMN21899597	3925057	95.0	3.6	resistant	99	0	0
H1_MAG_00034	SAMN21899598	1953820	94.2	3.6	sensitive	38	0	1
H1_MAG_00035	SAMN21899599	3406840	92.8	2.2	unclassified	5	0	0
H1_MAG_00036	SAMN21899600	2880608	95.0	5.0	resistant	18	0	0
H1_MAG_00037	SAMN21899601	7487990	95.7	5.8	resistant	33	0	0
H1_MAG_00038	SAMN21899602	5421061	93.5	3.6	resistant	8	0	0
H1_MAG_00039	SAMN21899603	9391796	97.1	7.9	resistant	312	0	2
H1_MAG_00040	SAMN21899604	5462335	92.1	3.6	resistant	149	0	0
H1_MAG_00043	SAMN21899605	3944493	89.9	2.2	unclassified	290	0	0
H1_MAG_00044	SAMN21899606	5489466	92.1	5.8	resistant	30	0	1
H1_MAG_00045	SAMN21899607	4972446	89.2	2.9	resistant	260	1	1
H1_MAG_00046	SAMN21899608	2981935	85.6	0.7	sensitive	7	0	0
H1_MAG_00047	SAMN21899609	2080170	86.3	1.4	resistant	27	0	0
H1_MAG_00048	SAMN21899610	5874343	85.6	1.4	resistant	59	1	4
H1_MAG_00049	SAMN21899611	4962459	92.8	9.4	resistant	86	1	1
H1_MAG_00050	SAMN21899612	3592967	82.0	0.0	unclassified	70	0	0
H1_MAG_00051	SAMN21899613	994875	81.3	0.0	unclassified	0	0	0

H1_MAG_00052	SAMN21899614	4182733	84.2	4.3	sensitive	390	0	1
H1_MAG_00053	SAMN21899615	2603677	82.7	3.6	resistant	331	0	0
H1_MAG_00054	SAMN21899616	829920	78.4	2.9	unclassified	15	0	0
H1_MAG_00056	SAMN21899617	2810691	76.3	0.7	unclassified	20	0	0
H1_MAG_00058	SAMN21899618	6268261	79.9	4.3	sensitive	46	0	0
H1_MAG_00059	SAMN21899619	2785643	78.4	2.9	unclassified	123	0	0
H1_MAG_00060	SAMN21899620	676081	74.8	0.0	unclassified	1	0	0
H1_MAG_00061	SAMN21899621	3518011	81.3	6.5	unclassified	46	0	0
H1_MAG_00062	SAMN21899622	2820293	75.5	2.9	sensitive	5	0	0
H1_MAG_00063	SAMN21899623	2111507	74.8	2.9	unclassified	11	0	1
H1_MAG_00064	SAMN21899624	3523699	74.8	4.3	resistant	58	0	1
H1_MAG_00065	SAMN21899625	2936184	77.0	7.2	sensitive	141	0	0
H1_MAG_00066	SAMN21899626	1213741	71.9	3.6	sensitive	14	0	0
H1_MAG_00069	SAMN21899627	3161246	71.9	7.9	resistant	76	0	1

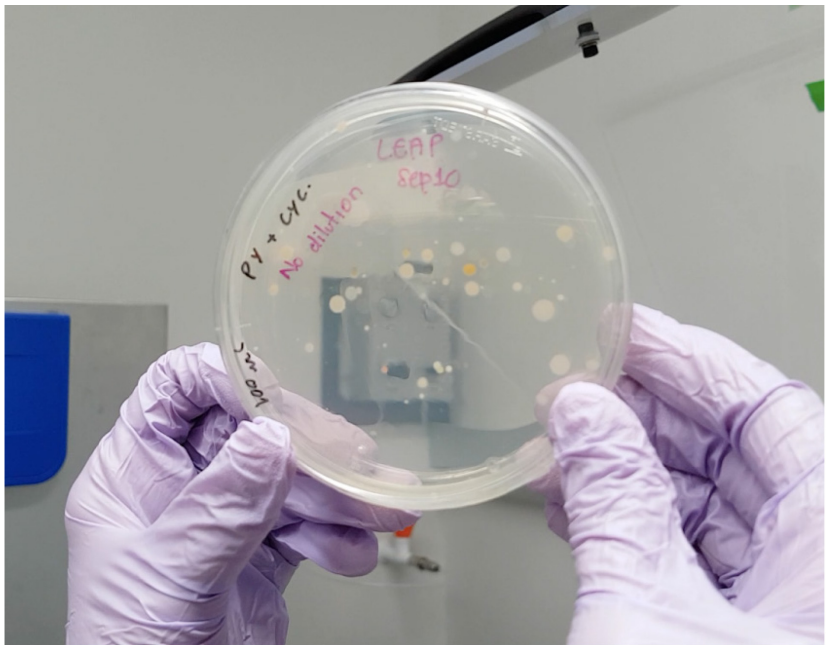
Table III.S4 Multiple linear regression model and variance partitioning of MAGs abundance in Phase II in control mesocosms. P-values are reported for each predictor, asterisks indicate significant *p*-values after Bonferroni correction (p<0.0125) and reports of significant factors are highlighted in bold. Adjusted R-squared equals 43.2 % for MAG abundance in controls as response variable (n=425, F-statistic: 78.7).

Response variable	Predictor	Estimate (SE)	t value	*p*-value	Explained variance
MAG mean abundance in Phase II control mesocosms	EPSPS classification :				2%
	- Sensitive	0.111 (±0.116)	0.95	0.343	
	- Resistant	**0.409 (±0.120)**	**3.41**	**0.001***	
	MAG antibiotic resistance potential	0.111 (±0.049)	2.26	0.025	1%
	MAG mean abundance in Phase I control mesocosms (log10)	**0.756 (±0.044)**	**17.14**	**<0.001***	40%
					Residuals: 58%

CHAPTER IV : Genome-wide selective sweeps rarely explain the ecological success of bacterial populations responding to a novel environmental stress

Naíla Barbosa da Costa[1,2], Olga Maria Pérez-Carrascal[3], Marie-Pier Hébert[2,4], Vincent Fugère[2,5,6], Andrew Gonzalez[4,6], Gregor Fussmann[2,4], B. Jesse Shapiro[1,2,7,8]

[1] Département des sciences biologiques, Université de Montréal, Montreal, Canada; [2] Groupe de Recherche Interuniversitaire en Limnologie et environnement aquatique (GRIL), Montreal, Canada; [3] Department of Integrative Biology, University of California, Berkeley, Berkeley, CA, USA; [4] Department of Biology, McGill University, Montreal, Canada; [5] Département des sciences de l'environnement, Université du Québec à Trois-Rivières, Trois-Rivières, Canada; [6] Québec Centre for Biodiversity Science (QCBS), Montreal, Canada; [7] Department of Microbiology and Immunology, McGill University, Montreal, Canada; [8] McGill Genome Centre, McGill University, Montreal, Canada



In prep.

Abstract

Evolutionary changes depend on shifts of heritable traits within a population and are traditionally considered slow compared to ecological changes, related to shifts in species abundance and composition within a community. Nevertheless, ecological and evolutionary processes may overlap, particularly in bacteria with short generation times and large population sizes. Here we asked to what extent within-species evolution coincided with ecological responses of bacterioplankton to pesticide contamination. We hypothesized pesticide sensitive species that are ecologically successful (i.e. increase in relative abundance after a pesticide pulse) experience adaptive evolution causing genome-wide selective sweeps that purge genetic diversity within-species. We measured evolutionary changes by tracking temporal changes of single nucleotide variants (SNVs) in bacterial populations from an 8-week experiment in freshwater mesocosms treated with pulses of a glyphosate-based herbicide (GBH). We observed populations with variable evolutionary dynamics, but with similar ecological trends in their responses to the GBH treatment, including one potential genome-wide selective sweep in *Aquidulcibacter*, found to be putatively sensitive to glyphosate based on the classification of the glyphosate target gene. The other ecologically successful populations are potentially resistant to GBH which may explain the observed lack of directional variation in intraspecific diversity and the absence of genome-wide selective sweeps, as the pesticide does not represent a selective pressure to them as it does to the sensitive species. This work provides experimental evidence of the stable ecotype model in a semi-natural environment. Moreover, it highlights how ecological success in the face of changing environments is rarely explained by genome-wide sweeps unless species are sensitive to the selective pressure in action. Ongoing investigations of a reproduction of this work will advance the analysis and validate conclusions using a replicated experimental design.

Introduction

Understanding how bacteria evolve and diverge into distinct groups in nature is challenging – arguably more so than for eukaryotes (Fraser, Alm, Polz, Spratt, & Hanage, 2009). In contrast with eukaryotes, prokaryotes reproduce strictly asexually and gene flow happens through horizontal gene transfer, which can transcend species boundaries and whose rate may vary from one lineage to another (Shapiro & Polz, 2014). Nevertheless, bacterial species may be conceived as cohesive units whose diversity is constrained by the forces of natural selection and recombination (Cohan, 2016, 2019; Shapiro, Leducq, & Mallet, 2016). In practice, species may be defined based on barriers to gene flow (Bobay & Ochman, 2017) or on a threshold of average nucleotide identity, usually >95% (Jain et al., 2018).

The stable ecotype model (Cohan, 2001; Cohan & Perry, 2007) is a prominent evolutionary model explaining cohesion within bacterial species that has received some support from observational data (Bendall et al., 2016; Jain et al., 2018). The model proposes that diversity within a population occupying a given ecological niche (i.e. an ecotype) is occasionally purged by periodic selection. According to the ecotype model, when an individual acquires an adaptive mutation, this mutant and its clonal descendants will outcompete others as a consequence of their increased fitness (a process also called clonal expansion) and replace the rest of the genetic diversity within that ecotype (Cohan, 2001). In other words, when recombination is relatively low, any genetic variation in the population is eventually purged by selection, resulting in a genome-wide selective sweep.

The stable ecotype model provides a mechanistic explanation for the origin of phylogenetically distinct groups based on genome-wide selective sweeps, which occur as divergent ecotypes undergo a series of clonal expansions driven by selection. Alternatively, gene-specific selective sweeps may occur within a population when recombination rates are relatively high. In this case, individual adaptive genes, rather than complete genomes, sweep to fixation (Shapiro, 2016). Bendall et al. (2016) were pioneers in detecting a genome-wide sweep in a natural bacterial population through a 9-year metagenomic study of a freshwater lake. By tracking variations in single-nucleotide polymorphisms and gene frequencies in 30 reconstructed bacterial genomes, they detected one genome-wide diversity purge within a population of *Chlorobium*-111.

Additionally, the authors inferred that gene-specific sweeps occurred before the beginning of the study in other six populations, that had loci with unexpectedly low diversity compared to the genome-wide average. However, Bendall et al. (2016) did not measure any selective pressure and thus they could not formally conclude if the genome-wide and gene-specific sweeps were caused by selection – predicted to be the cohesive force constraining within-species diversity in the stable-ecotype model – as opposed to random demographic changes leading to genetic drift.

Prokaryotes are commonly faced with changing selective pressures, such as phage predation (Rodriguez-Valera et al., 2009), nutrient limitation (Allison & Martiny, 2008), temperature variation, and the presence of contaminants (Coleman & Chisholm, 2010; Shade et al., 2012). These environmental perturbations may affect ecological processes, such as community stability (Shade et al., 2012), and evolutionary processes, such as changes in heritable traits within a population (Hutchins et al., 2015). Although evolutionary processes are traditionally considered slow compared to ecological processes, evolution can actually be explained by the conjunction of ecology and genetics, as it involves processes such as birth and death, competition, mutation and selection (Lenski, 2017). Both ecological and evolutionary timescales may overlap in prokaryotic species due to their typically short generation times and large populations.

We previously showed how ecological processes and the composition of genes changed due to the presence of a glyphosate-based herbicide (GBH) within experimental communities (Barbosa da Costa et al., 2021, 2022). Bacterial community composition shifted in response to the herbicide treatment (Barbosa da Costa et al., 2021), which also altered the composition of genes, cross-selecting for antibiotic resistance genes in the community (Barbosa da Costa et al., 2022). The relative abundance of genomes present after a high dose of GBH was better predicted by the number of antibiotic resistance genes they encoded than by known resistance mutations in the glyphosate target enzyme (Barbosa da Costa et al., 2022).

Here, we used a controlled experimental setup with defined selective pressures to investigate if bacterial populations with similar ecological responses would experience the same evolutionary responses in the semi-natural conditions of artificial ponds filled with pristine lake water. We hypothesized that selective sweeps would occur over relatively

short time scales (8 weeks) in the presence of a strong selective pressure, due to bacterial rapid adaptation to the stressor. Specifically, we expect that ecologically successful species without an evident resistance mechanism against the stressor will experience genome-wide selective sweeps that purge genetic diversity across their genomes.

To test this hypothesis, we exposed bacterial communities to a GBH in replicated 1,000L freshwater mesocosms, imposing the selective pressure of pulse perturbations promoted by agrochemical contamination. From shotgun metagenomic sequencing, we identified 11 populations with sufficient coverage to track within-species single nucleotide variants (SNVs) over time and across experimental treatments. While allele frequencies remained approximately stable in six populations, diversity decreased slightly in three populations and increased in two populations after GBH pulses. We detected a possible genome-wide selective sweep in one population of the initially glyphosate-sensitive *Aquidulcibacter*, whose diversity dropped from about 3,000 to 121 SNVs/Mbp after the final and strongest GBH pulse. In summary, varied and idiosyncratic evolutionary dynamics accompany rapid ecological responses of bacterioplankton to environmental stressors, but we highlight that the populational dynamics in the two potentially glyphosate-sensitive species found in our study confirm the expected trend towards reduced intraspecific variation when facing strong stress. A potential genome-wide selective sweep was detected in the only glyphosate-sensitive population among ecological successful species. A future reproduction of this study will help confirm if this trend was the result of a random or directional evolutionary process.

Methods

Experimental design and sample collection

To control selective pressures in a semi-natural setting, we performed an eight-week mesocosm experiment at the Large Experiment of Ponds (LEAP) in 2016 from August 17th (day 1) to October 12th (day 57), as previously described (Fugère et al., 2020). The LEAP platform is located at McGill University's Gault Nature Reserve (QC, Canada) and eight mesocosm ponds were filled with 1,000 L of water from a pristine upstream lake (Lake Hertel 45°32' N, 73°09' W). The water was sieved to prevent the introduction of fish and large debris while allowing planktonic communities to pass

through the coarse mesh. Previous studies have described the responses of bacterioplankton (Barbosa da Costa et al., 2021, 2022), zooplankton (M. P. Hébert et al., 2021) and phytoplankton (Fugère et al., 2020) to the experimental treatments that here are explored through the lens of population genomics in bacterial populations.

We applied pulses of the GBH Roundup® Super Concentrate Grass and Weed Control (Bayer ©; reg. #22759) to mesocosm ponds and sampled for bacterial DNA in 10 timepoints from experiment day 1 to 57, as illustrated in Figure IV.1. A commercial formulation was chosen as it is one of the many pesticide formulations available to use in contrast with the pure glyphosate salt that is not applied to fields and would be unaffordable in such a large-scale field experiment. The eight experimental ponds here studied are a subset of treatments selected for metagenomic sequencing analyses, the full set of treatments is described in Hébert (2021) and Barbosa da Costa (2021).

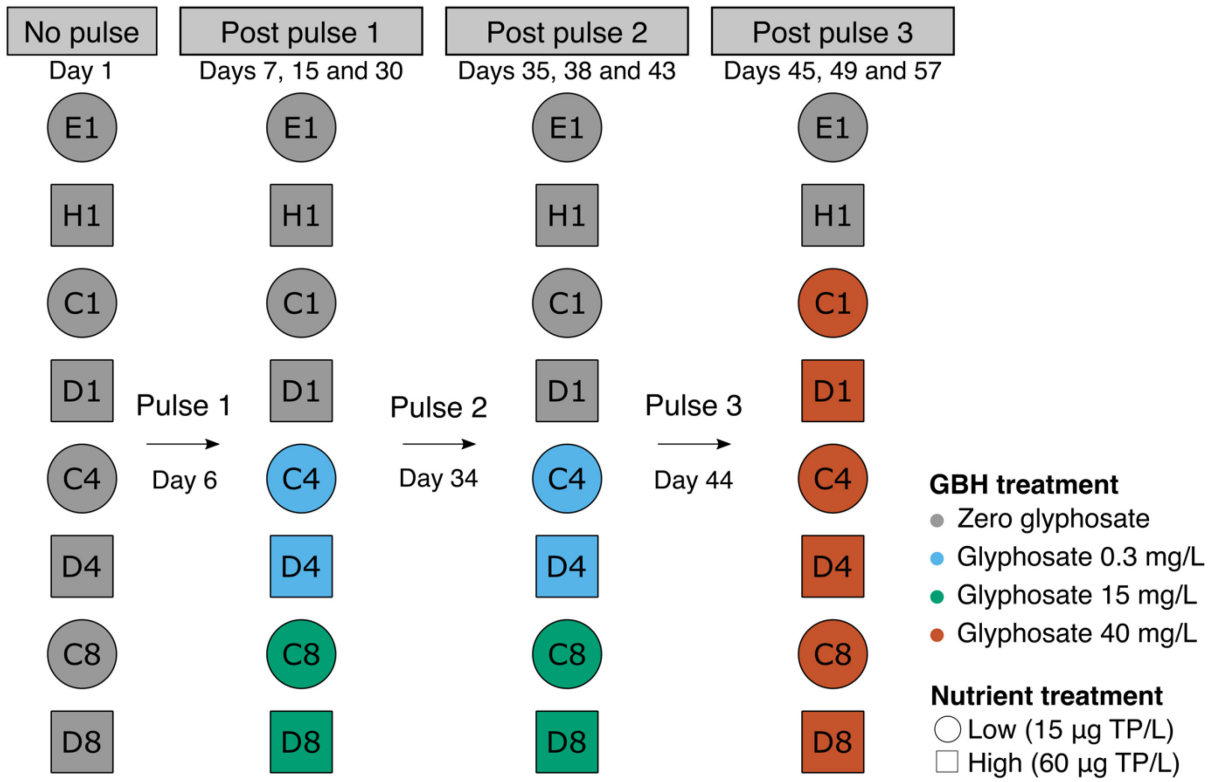


Figure IV.1 Schematic representation of the experimental design. On day 1, before the application of pulse 1, all ponds had zero measurable glyphosate. Between days 7 and 43, half of the ponds received two pulses of a glyphosate-based herbicide (GBH) treatment in different concentrations and the other half were kept without any pesticide.

After day 44, ponds previously treated with GBH and two controls received a higher dose of glyphosate while other two were left intact. The same GBH treatment was replicated in ponds with low or high nutrient treatment.

Pulses of the GBH were applied on days 6, 34, and 44 of the experiment. The first two pulses varied according to the treatment to reach target values of glyphosate (acid equivalent) of 0.3 mg/L (ponds C4/D4) and 15 mg/L (ponds C8/D8). Other ponds (C1/D1 and E1/H1) did not receive any glyphosate input at these two pulses (Figure IV.1). The last pulse had a higher dose (40 mg/L glyphosate) and was applied to all ponds except E1/H1. Half of the ponds received a nutrient treatment containing phosphate (KH_2PO_4 and K_2PO_4) and nitrate (KNO_3) to reach 15 µg P/L and 231 µg N/L, creating a mesotrophic (low-nutrient) environment while the other half received the same nutrients for target concentrations of 60 µg P/L and 924 µg N/L to create a eutrophic (high-nutrient) environment. These nutrient sources were added in combination to keep the lake N:P molar ratio of 33.

We collected water samples in ten timepoints: day 1, 7, 15, 30, 35, 38, 43, 45, 49 and 57, as noted in the top of Figure IV.1. We used integrated PVC tubing samplers (35 cm long, 2.5 cm diameter) and filtered 250 mL through Millipore hydrophilic polyethersulfone membranes of 0.22 µm pore size (Sigma-Aldrich, St. Louis, USA). Filters were stored at -80 ºC until DNA extraction.

DNA extraction, metagenomic sequencing and sequence preprocessing

We extracted DNA from the 80 samples (8 ponds in 10 timepoints each) used in this study with the PowerWater DNA Isolation kit (MoBio Technologies Inc.) following the manufacturer's guidelines. We included in the DNA isolation protocol a step to enhance cell lysis, consisting of a 5-min vortex agitation of the filter with beads in the lysis buffer.

We performed shotgun metagenomic paired-end sequencing through Illumina HiSeq 4000 technology (2 x 100 bp reads). Libraries had 50 ng of genomic DNA from each sample and were prepared with the NEBNext Ultra II DNA Library Prep kit for Illumina (New England Biolabs®) as per the manufacturer's recommendations (390 bp average fragment size).

We used FastQC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/) to check the overall sequencing quality of fastq read files after the removal of Illumina

adapters and quality filtering with Trimmomatic (Bolger et al., 2014) in the paired-end mode. The average number of reads per sample after quality filtering was 32.5 million reads (min 19.1 million max 73.4 million).

Metagenomic de novo coassembly, binning, dereplication and curation of MAGs

We co-assembled reads from each of the eight experimental ponds separately (Figure IV.1) using MEGAHIT v1.1.1 (D. Li et al., 2015), as described in Barbosa da Costa (2022). The dataset used for the co-assembly included one additional timepoint (day 41) that was not included in this study, as here we intended to keep a balanced number of samples after each pulse (*n*=3). Each set of 11 metagenomes from different timepoints corresponding to the same experimental pond was used in the co-assembly and we set a minimum contig length of 1 kbp.

We used Bowtie2 v2.4.2 (Langmead & Salzberg, 2012) within anvi'o v5.1 to map metagenomic reads to each set of scaffolds, then anvi'o to profile each BAM output thus estimating coverage and detection of scaffolds across timepoints within the same pond. Scaffolds were clustered into bins through the automatic binning algorithm CONCOCT (Alneberg et al., 2014) within anvi'o v5.1 (Eren et al., 2015). We estimated the completeness and redundancy of each bin using anvi'o with HMMER v3.2.1 (Eddy, 2011) to identify single-copy core genes belonging to bacterial or archaeal genomes.

Bins were manually refined and dereplicated as previously explained in Barbosa da Costa (2022) following the guidance of anvi'o developers (Delmont et al., 2018; Eren et al., 2015). Curated bins with a minimum 70% completeness and maximum 10% redundancy were classified as metagenome-assembled genomes (MAGs) and have been submitted to NCBI together with the metagenomic reads set under BioProject PRJNA767443.

Selection of MAGs for SNV profile and taxonomic assignment

To select the MAGs that responded most strongly to the GBH treatment in addition to the less responsive MAGs, we used Principal Response Curves (PRCs) performed with the abundance matrix of 426 non-redundant MAGs identified throughout the experiment, as described in Barbosa da Costa (2022). PRCs are used to infer the effect of experimental treatments on community composition across time. A PRC model returns species scores that quantify how much each species contributed to the overall change in

community structure (Auber et al., 2017; Van den Brink et al., 2009; Van den Brink & Ter Braak, 1999). In other words, the species scores are proportional to how much a species responded to the treatment. We thus selected 31 MAGs for Single Nucleotide Variant (SNV) profiling based on their species score: the top 16 most responsive MAGs, positively affected by GBH treatments (score > 0.07) and the 15 least responsive MAGs (-0.007 < score < 0.007). The 31 selected MAGs are summarized in Table IV.S1, which reports their respective NCBI accession number, genome size, taxonomic classification, PRC score, putative classification of the EPSPS enzyme detected in their genome and the inferred number of Antimicrobial Resistance Genes (ARGs) estimated as shown Barbosa da Costa et al. (2022) in using the Resistance Gene Identifier (RGI) application and the Comprehensive Antibiotic Resistance Database (CARD) database. To complement the previous taxonomy assignment of CheckM (Parks et al., 2015) performed with anvi'o, which returned some low-resolution classifications at the phylum or class level, we now include the classification of these 31 selected MAGs through the Genome Taxonomy Database Toolkit (GTDB-Tk) (Chaumeil, Mussig, Hugenholtz, & Parks, 2020).

Detection of intraspecific diversity in populations using MAGs as references

To compare intraspecific diversity of populations present within the same treatment pond after different GBH pulses, we inferred SNVs within samples using as reference genomes the 31 selected MAGs which were less- or high-responsive to treatments (Table IV.S1). Reads were mapped with Bowtie2 v2.4.2 (Langmead & Salzberg, 2012) and, to increase overall coverage depth, we merged alignments of timepoints that followed the same pulse using the "samtools merge" function within SAMtools v1.13 (H. Li et al., 2009). To ensure high coverage of reads mapped to the reference genome, 3 timepoints were merged by pulse, as indicated in Figure IV.1: days 7, 15 and 30 were grouped post pulse 1; days 35, 38 and 43 post pulse 2; and days 45, 49 and 57 post pulse 3. As only one timepoint was sampled before any pulse application, the low coverage (generally <5X) of the selected MAGs prevented us from inferring SNVs on experiment day 1.

To measure genetic variation within the population represented by each MAG, we used inStrain v1.5.7 (Olm et al., 2021). We ran the "inStrain profile" function to call SNVs from alignments with a minimum MAPQ of 13. We also set a minimum coverage depth of 5X to call a variant at a nucleotide position and a 99% minimum percent identity of read

pairs to consensus. At this high level of nucleotide identity, we are very likely to be sampling diversity within rather than between species. SNVs were inferred strictly from paired reads and we used the function “inStrain compare” to differentiate positions where the reference frequency was 1 from positions that did not pass the aforementioned requirements, which were then identified as “NA” while comparing a single SNV among different samples. We filtered out SNVs found within 100 bp of scaffold extremities and positions with very low or very high coverage values compared to the median position coverage within the sample (each sample corresponds to a MAG found in a specific pond after a specific pulse), as variants found in these positions could be a result of sequencing bias, mapping, or assembly errors. We kept only positions whose coverage was greater than 30% of the median coverage and smaller than 3 times the median. We used Prodigal v2.6.3 (Doug Hyatt, Gwo-Liang Chen, Philip F LoCascio, Miriam L Land, Frank W Larimer, 2010) to predict genes in reference genomes, and inStrain to classify mutations as intergenic, synonymous or non-synonymous. In rare cases, SNVs remained unclassified because the algorithm was unable to decide whether the mutation was synonymous or non-synonymous.

Results

Tracking SNV frequency changes within MAG populations over time

After metagenomic sequencing from our experimental ponds, we assembled MAGs and identified the 16 species that responded most positively to the GBH treatment and the 15 less affected species (Table IV.S1). Of these 31 MAGs, 11 were present with sufficient coverage after all 3 pulses in treatment ponds to track SNV frequencies over time and in response to treatments (Figure IV.S1). The metagenomic reads mapping to the same MAG in the same pond are hereafter referred to as populations, as they might belong to individuals of the same species found in the same location and that can interchange genetic material. The genome coverage of the reference MAG within these populations was on average 21X (min 1X max 103X, Table IV.S2) and the breadth of coverage averaged 89% (min 47% max 100%, Table IV.S2).

The detection of SNVs increased with greater genome coverage within samples (Figure IV.S2, Linear Model R^2=0.31, t=4.6, p<0.001). This could bias our analysis and

thus requires a careful investigation when comparing SNV counts over time or across treatments within populations. However, high SNV counts were not necessarily associated with deeper coverage in some MAGs. For example, in samples of the pond C8, the MAG classified as *Nevskia* 1 (C8_MAG_00031, Table IV.S1) had coverage of 7X after pulse 1 and 68X after pulse 2 and in both timepoints it had similarly high SNV density (respectively 1386 and 1654 SNVs/Mbp, Table IV.S2). Similarly, *Aquidulcibacter* (C4_MAG_00010, Table IV.S1) in pond D8 had a genome coverage of 45X after pulse 1 and 103X after pulse 3, while the number of SNVs/Mbp were respectively 2943 and 3277 (Table IV.S2). We conclude that SNV changes cannot be attributed to variation in sequencing coverage alone, but we are careful to consider this potential confounder in the results that follow.

Of the 15 populations shown in Figures IV.S1 and IV.S2 (i.e. distinct MAGs in different ponds), four exhibited SNV variation proportional to MAG coverage and we thus did not explore how they changed in intraspecific diversity through time (Table IV.S2). For each of the remaining 11 populations, we plotted changes in two measures of genetic diversity over time and GBH pulses: the median minor allele frequency (MAF) and the density of SNVs in the genome. These measures revealed distinct evolutionary dynamics in each of the different populations (Figure IV.2). As a measure of natural selection acting at the protein level, we classified each SNV as synonymous (S), non-synonymous (N), or intergenic. Most populations had relatively stable and low N:S ratios over time and GBH pulses (Figure IV.S3). In four populations, however, the N:S ratio was higher and more variable across pulses (Figure IV.S3A), and these are also the populations with the lowest densities of SNVs (average 40 SNVs/Mbp, min 6 max 101). Such low rates of genetic diversity and high N:S ratios suggest a relatively small effective population size and/or recent population bottleneck (Rocha et al., 2006).

Intra-specific diversity increases, decreases, or remains stable after GBH pulses in different populations

Populations varied in terms of their evolutionary responses (i.e. SNV dynamics) to GBH pulses. Four populations (Figure IV.S4) could not be compared from one pulse to another because their MAG coverage was proportional to the SNV density, preventing an accurate interpretation of the results. Of the remaining populations, six maintained a

roughly stable median MAF (Figure IV.A-F; Figure IV.3), while three showed a decrease (Figure IV.G-I; Figure VI.4A-C) and two showed an increase (Figure IV.J-K; Figure VI.4D-E) in median MAF after one or more pulses.

Of the six populations with relatively stable diversity, four were the populations with the lowest SNVs/Mbp and lowest MAF (Figure IV.2A-D), as well as the highest N:S ratios (Figure IV.S3A). These four MAGs are all classified as encoding the putatively glyphosate-resistant EPSPS gene (Table IV.S1). Three of them had a high proportion of fixed alleles (Figure IV.3A-C), consistent with a recent population bottleneck before pulse 1.

The two other populations with steady diversity correspond to the same MAG, the alphaproteobacterium UBA2784 with the resistant EPSPS gene (Table IV.S1), found in the high nutrient pond D8 (Figure IV.3E) and in the low nutrient pond C8 (Figure IV.3F). In both ponds, UBA2784 maintained a relatively high and consistent diversity in terms of SNVs/Mbp and median MAF (Figure IV.E-F) with relatively few fixed SNVs dominated by a single allele (Figure IV.3E-F).

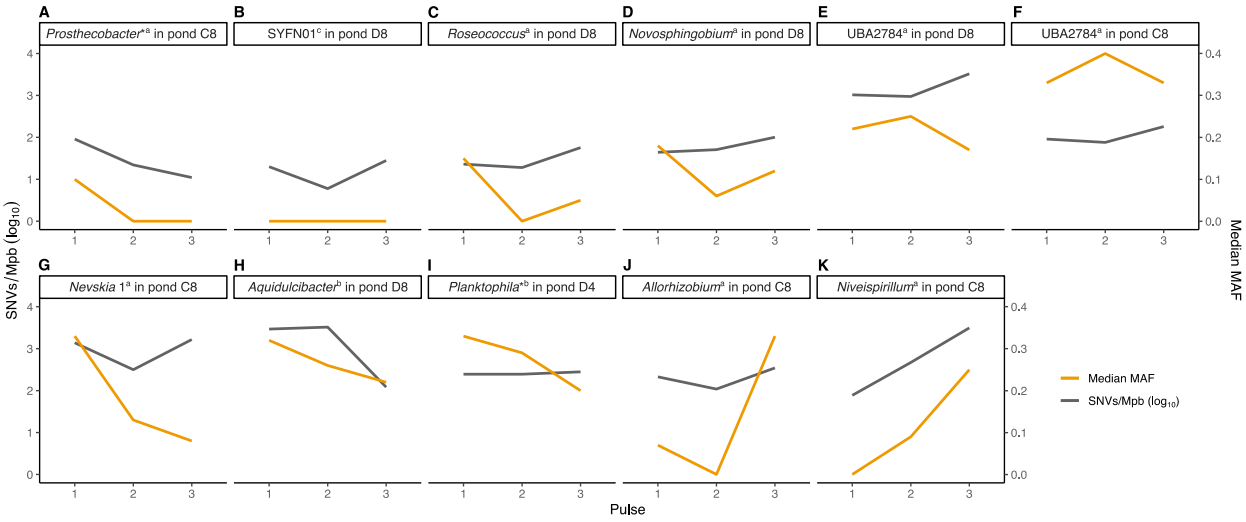


Figure IV.2 Variable evolutionary dynamics across bacterial populations. Six populations (A-F) showed relatively stable median minor allele frequencies (MAFs) across the three pulses, three populations decreased in median MAF (G-I), and two increased (J-K). *MAG with PRC score close to zero; all others have high positive scores, indicating an increase in relative abundance in the presence of GBH. [a]Resistant EPSPS gene found in the genome, [b]Sensitive EPSPS gene found in the genome, [c]Unclassified EPSPS gene.

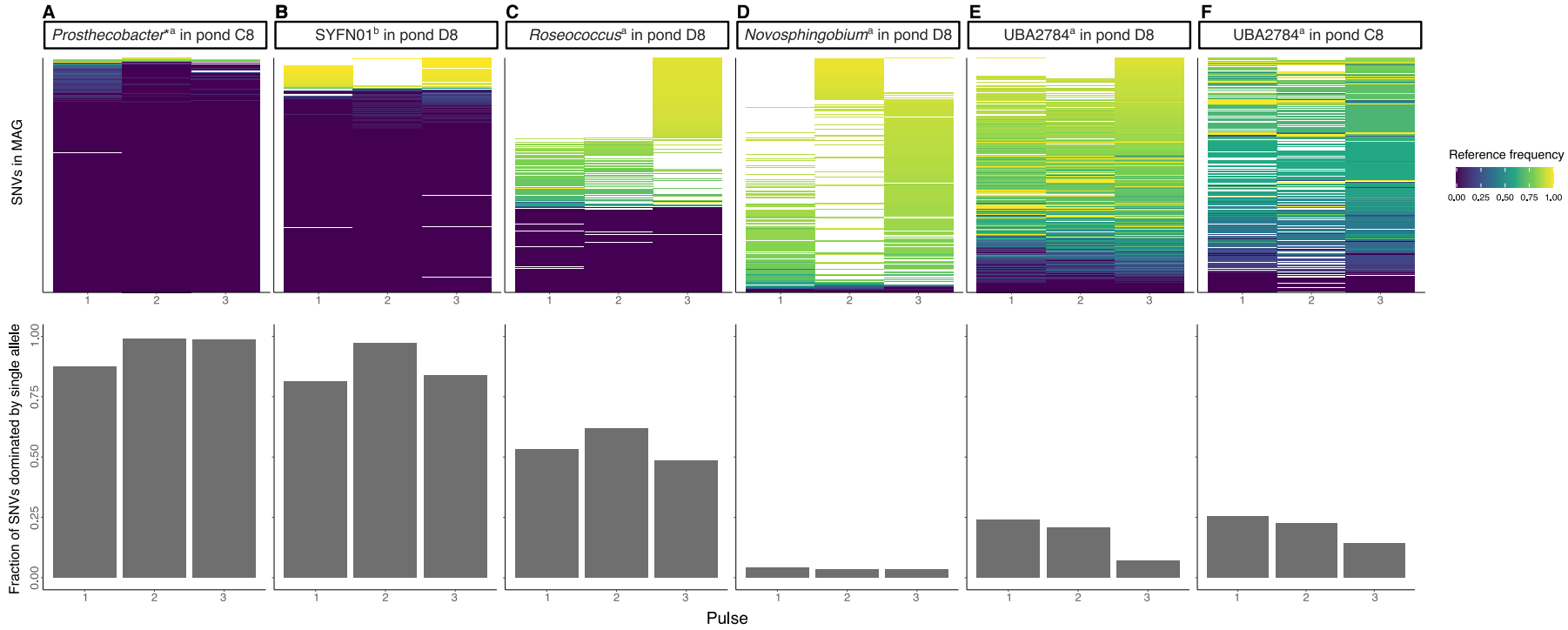


Figure IV.3 Evolutionary dynamics of populations with relatively stable allele frequencies. Top panels (heatmaps) show the reference allele frequency at SNV positions with at least 5X coverage. The bottom panel shows the fraction of SNV positions dominated by a single allele. *MAG with PRC score close to zero, [a]Resistant EPSPS gene found in the genome, [b]Unclassified EPSPS gene

Three populations decreased in diversity after pulse 2 or 3 (Figure IV.2G-I). They began with high median MAFs (>0.25) in pulse 1 but over time the frequency of the major allele increased, approaching fixation and pushing the median MAF below 0.25. Two of these MAGs had high PRC scores (>0.07), meaning that they were favoured by the GBH treatment, while the third, *Planktophila* (D4_MAG_00049, with the sensitive EPSPS gene, Table IV.S1) had a PRC score close to zero, meaning it was less responsive to the GBH treatment. The decline in *Planktophila* genetic diversity was relatively subtle (Figure IV.2I) and the fraction of fixed SNVs remained relatively low across pulses (Figure IV.4C). These observations are inconsistent with a genome-wide sweep in this population which persisted, but did not particularly thrive in the presence of GBH stress.

The two populations with more marked decreases in diversity and high PRC scores were a *Nevskia* 1 (C8_MAG_00031, with the resistant EPSPS gene, Table IV.S1) found in pond C8 and an *Aquidulcibacter* (C4_MAG_00010, with a sensitive EPSPS gene, Table IV.S1) found in pond D8. While *Nevskia* 1 had a high MAF only during pulse 1 and a drop in SNV density during pulse 2 (Figure IV.2G) – which cannot be explained by MAG coverage (Table IV.S2) – *Aquidulcibacter* experienced a drastic reduction in SNV density in pulse 3 (Figure IV.2H) in what could the beginning of a genome-wide sweep, considering that many alleles were fixed (Figure IV.4B). In pulses 1 and 2, the *Aquidulcibacter* had about 3000 SNVs/Mbp and after pulse 3 it dropped to 121 SNVs/Mbp (Table IV.S2). Although there was also a reduction in MAG coverage from 45X and 103X in pulses 1 and 2 to 6X in pulse 3, we see a trend towards allele fixation in pulse 3 at genomic positions previously identified with variability (Figure IV.4B). The relatively low coverage of 6X in pulse 3 would have the tendency to push estimated allele frequencies closer to the extremes (0 or 1) by chance. However, the observation that most of the SNVs in *Aquidulcibacer* in pulse 3 were close to a reference allele frequency of 1 (visible as the block of yellow in the heatmap; Figure IV.4B), with a near-absence of zero frequencies, is consistent with a selective sweep or bottleneck rather than a sampling artifact.

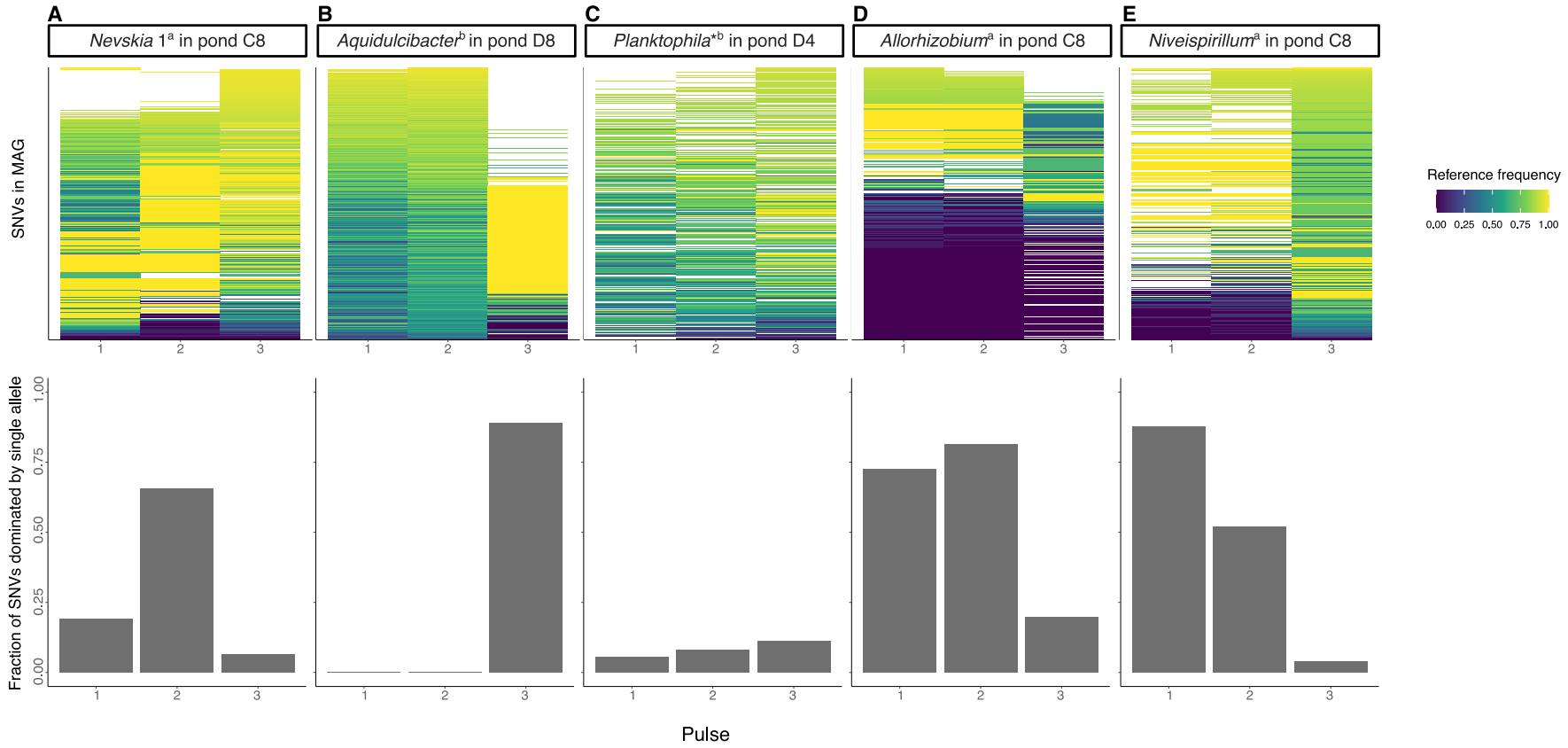


Figure IV.4 Evolutionary dynamics of populations with directional changes in allele frequencies. Top panels (heatmaps) show the reference allele frequency at SNV positions with at least 5X coverage. The bottom panel shows the fraction of SNV positions dominated by a single allele. Panels A-C show populations with declining genetic diversity over time and D-E with increasing diversity. *MAG with PRC score close to zero, [a]Resistant EPSPS gene found in the genome, [b]Sensitive EPSPS gene found in the genome

Two populations increased intraspecific diversity after pulse 3: an *Allorhizobium* and a *Niveispirillum* found in pond C8. Their median MAF increased from 0-0.07 or 0.09 in the first two pulses to 0.33 or 0.25 in pulse 3 (Figure IV.2J-K). The density of SNVs increased in both populations, mainly in *Niveispirillum* (Figure IV.2K), and cannot be explained by variation in MAG coverage (Table IV.S2). Both MAGs have the putatively resistant EPSPS gene (Table IV.S1) and had many genomic positions with fixed alleles in pulses 1 and 2, giving way to fewer fixed alleles and more diversity in pulse 3 (Figure IV.4D-E).

Discussion

We used an experimental freshwater system to assess how bacterial community responses to a common anthropogenic stressor are accompanied by evolutionary changes in SNV diversity within bacterial populations. We show that species responding to a GBH treatment similarly within the community vary in their response at the population level as they exhibit divergent trends in allele frequency dynamics. Differential evolutionary responses regarding changes in intraspecific diversity may be explained by recent evolutionary history (Shapiro, 2016), by the evolution of glyphosate resistance (Rainio et al., 2021) or by unmeasured selective pressures within the community such as phage predation (Cordero & Polz, 2014; Rodriguez-Valera et al., 2009).

The pioneering study of Bendall (2016) that detected a genome-wide sweep in nature over a 9-year time-course was re-analyzed by Shapiro (2016), who classified the 30 bacterial genomes into "old and diverse" populations or "young and low-diversity" populations. Shapiro (2016) concluded that old and diverse populations tended to maintain stable genetic diversity over the years while most of the young and low-diversity populations with a history of genome-wide sweeps would be more prone to experience further sweeps – which is more readily explained by small population sizes and drift than by selective sweeps. In our experiment, we applied a known selective pressure (GBH pulses), allowing us in principle to distinguish purges of genetic diversity due to drift from those due to selective sweeps.

We identified four populations with stable, low intraspecific diversity, high N:S ratio, and low SNV density, of which three had a putative glyphosate-resistant EPSPS gene,

and the other was unclassified. With an average of 40 SNVs/Mbp and median MAF below 0.2, they are potentially clonal populations of glyphosate-resistant species, which could explain the positive effect of the GBH treatment on the relative abundance of most of them (denoted by the high PRC score). The high N:S ratio suggests that these are young populations (Shapiro, 2016), possibly recently recovered from a population bottleneck or genome-wide sweep, after which purifying selection has had limited time to purge nonsynonymous mutations (which are more likely to be deleterious) from the population (Kryazhimskiy & Plotkin, 2008; Rocha et al., 2006). It is possible that these populations underwent a very rapid genome-wide selective sweep between the beginning of the experiment and the beginning of pulse 1, i.e. they were resistant to glyphosate since pulse 1 but not before it, but unfortunately, the MAGs lacked sufficient coverage before pulse 1 to test this hypothesis.

The alphaproteobacterium UBA2784 was the only MAG with stable high diversity and low N:S ratio, suggesting the maintenance of longstanding genetic diversity. It has a glyphosate-resistant EPSPS gene and was found in ponds receiving high amounts of GBH in different nutrient backgrounds. Although the median MAF was somewhat lower in the high nutrient treatment, it still retained a relatively high SNV density (~1700 SNVs/Mbp). The populations from both low- and high-nutrient treatments were far from being clonal, as the detected SNVs positions did not show alleles close to fixation. As these populations come from the same species and lake source they could be originating from a single old and diverse population that tends to retain genetic diversity through time, as observed by Shapiro (2016) in populations with low N:S ratio and high SNV density.

Three populations of different MAGs (identified as *Planktophila*, *Aquidulcibacter* and *Nevskia*) showed a decreasing MAF after a GBH pulse in the course of our experiment. *Planktophila*, one of the most ubiquitous and abundant freshwater bacterial genera (Mondav et al., 2020), was one of the few species not favoured by the GBH treatment that also had enough coverage for SNV profiling. Although expected to be abundant in the bacterioplankton, the MAG of *Planktophila* had the glyphosate-sensitive EPSPS class which could explain why it was not positively affected by the GBH treatment, and a declining population size could explain the reduction in intraspecific diversity after the last GBH pulse.

In contrast, *Aquidulcibacter* also had a sensitive EPSPS gene but nevertheless responded positively to the GBH treatment. It had a drastic decrease in diversity at the third GBH pulse that could be considered a genome-wide sweep since there was a large reduction in SNVs/Mbp and many alleles were fixed in previously variable positions. It is possible that the surviving clonal population had an alternative mechanism to resist glyphosate contamination that was selected by the last pulse, such as a degradation pathway (Hove-Jensen et al., 2014). *Aquidulcibacter* was also found in a control pond that never received a GBH pulse, enabling future comparative genomic analyses to determine the possible genetic targets of selection. Additionally, this experiment has been reproduced with replicates and ongoing analyses will corroborate to test the hypothesis that GBH drove the genome-wide sweep in the *Aquidulcibacter* population. Although genome-wide or gene-specific sweeps do not rely on replication to be confirmed in nature (Bendall et al., 2016), a replicated design with controls and treatments allows for testing the hypothesis of the sweeps being caused by a selective pressure.

The third population with decreased MAF after a GBH pulse was a MAG classified as the gammaproteobacterium *Nevskia*, whose genome has a glyphosate-resistant EPSPS. It is among the most positively affected species by the experimental treatment and the reduction in MAF after the second pulse was accompanied by an increase in the fraction of fixed alleles, which could indicate an incomplete sweep that perhaps extinguished sensitive variants. This reduction in intra-specific diversity was followed by an apparent recovery in the third pulse when SNV density was higher and fewer alleles were fixed. This MAG was also found in a pond that received only the last and highest GBH dose. Future gene and SNV analyses will allow testing for convergent evolution happening in populations of different environments where similar selective pressures were applied.

Two other populations of MAGs with the resistant EPSPS (*Allorhizobium* and *Niveispirillum*) increased in median MAF after pulse 3. This could be the result of a larger population size after the last GBH pulse, negative frequency-dependent or diversifying selection. "Kill-the-winner" dynamics has been invoked to explain the maintenance of high diversity lineages within a species through phage-predation (Rodriguez-Valera et al., 2009). An analysis of the diversity of CRISPR sequences in these populations would

provide information on their acquired viral resistance (Sorek, Kunin, & Hugenholtz, 2008) and clarify if this ecological interaction could explain why diversity increased in these populations. These two MAGs were also found after the last pulse in a pond that only received this highest GBH dose and future analyses comparing these populations would also allow testing for convergent evolution.

Future avenues of this work include quantifying parallel evolution, i.e. the independent emergence of similar genetic changes in different lineages (Tenaillon et al., 2016), in addition to detecting gene-specific sweeps, which should be more frequent than genome-wide sweeps in species with high recombination rates (Shapiro & Polz, 2014). Most populations in our study show no clear evidence for genome-wide selective sweeps, but gene-specific sweeps could have occurred.

Our study provides evidence that ecological changes occurring in a short 8-week time scale are accompanied by distinct and diverse evolutionary dynamics. These variable evolutionary responses are likely related to the fact that populations of resistant species (i.e. those with the resistance EPSPS class or with cross-resistance genes, such as ARGs) do not experience the same selective pressure as glyphosate-sensitive populations. Additionally, variable evolutionary responses may reflect the standing genetic diversity, different evolutionary histories, and ecological interactions of different populations within the community. Our study provides evidence of a potential genome-wide selective sweep in a semi-natural context confirming the stable ecotype evolutionary model that has implications for understanding the process of bacterial differentiation. Future analyses will confirm the causality of this sweep by comparing the population dynamics in treatments and controls and further explore the relative importance of other types of evolutionary processes in action, such as gene-specific sweep.

Funding: This study was supported by a Canada Research Chair and NSERC Discovery Grant to B.J.S. N.B.C. was funded by FRQNT and NSERC-CREATE/GRIL fellowships. LEAP was built and operated with funds from a CFI Leaders Opportunity Fund, NSERC Discovery Grant, and the Liber Ero Chair to A.G.

Acknowledgements: We are thankful to A. Arkilanian, C. Normandin, D. Maneli, and T. Jagadeesh for aid in the field and to J. Marleau for support in the laboratory.

Supplementary information

Supplementary figures

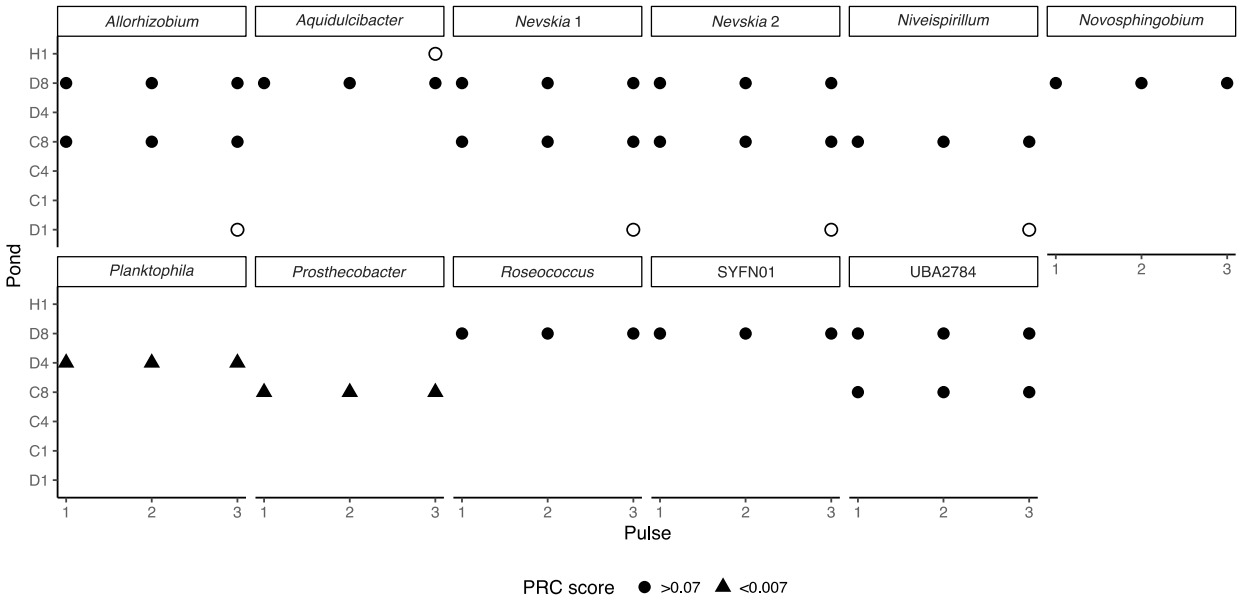


Figure IV.S1 Distribution of MAGs across ponds and temporal GBH pulses. Of the 53 samples with enough genome coverage for SNV calling showed here, we selected the 45 (filled shapes) comprising MAG populations (i.e. the same MAG in the same pond) found after pulses 1, 2 and 3 to compare SNV profiles over time. Only two out of the 12 MAGs with enough coverage to profile SNVs had low PRC scores (<0.007). Samples not included in this study (non-filled shapes) will be the subject of a future study on parallel evolution of populations of the same MAG in different environments (i.e. ponds/treatments). Note: because two different MAGs were identified as *Nevskia* (table IV.S1), one is here being called *Nevskia* 1 (C8_MAG_00031) and other *Nevskia* 2 (D8_MAG_00048).

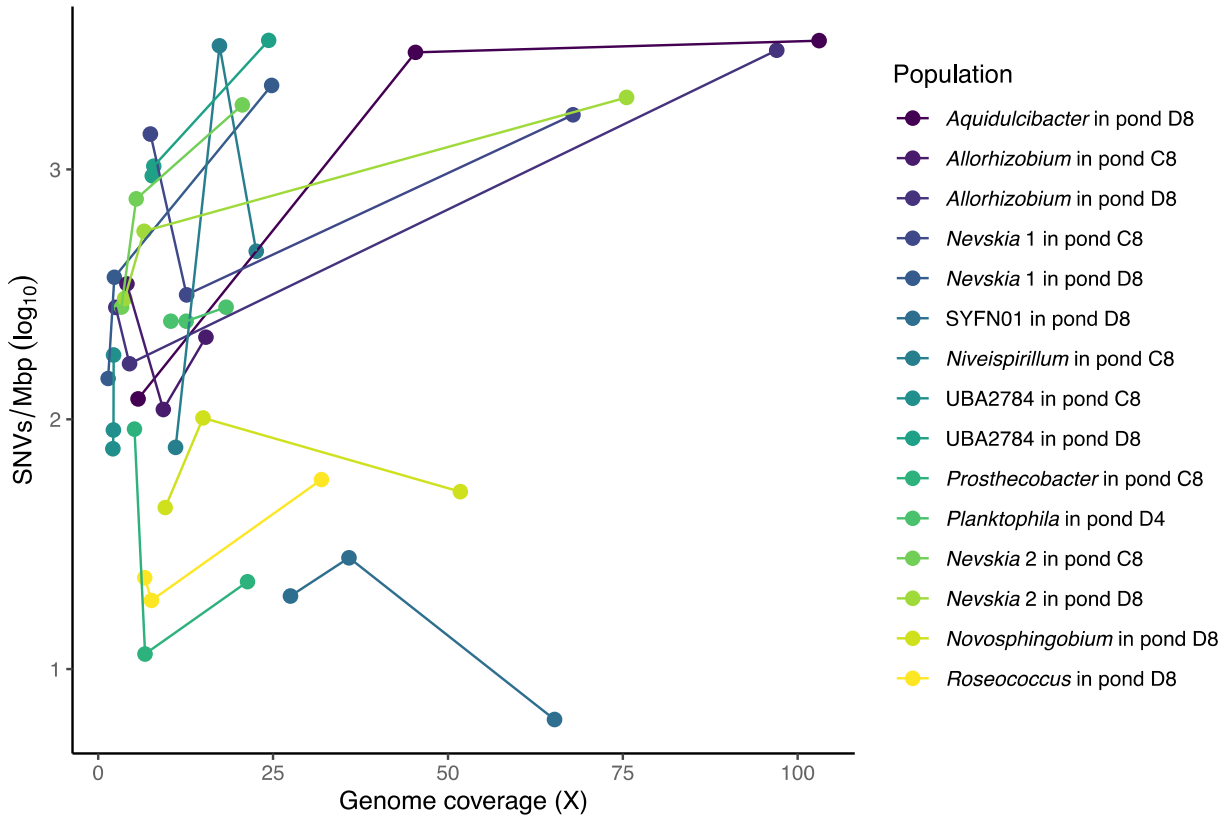


Figure IV.S2 Genome coverage within samples increases the detection of SNVs, although MAGs with high coverage do not necessarily have high SNVs/Mbp.

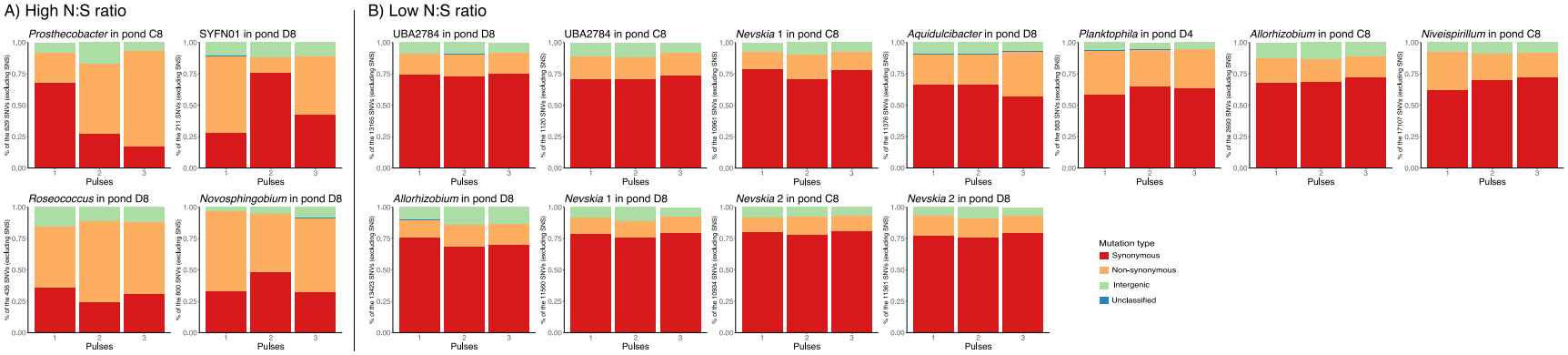


Figure IV.S3 Proportion of SNVs according to mutation type after pulse 1, 2 and 3 in populations of MAGs with A) high N:S and B) low N:S ratios (note: the four populations at the bottom are not analyzed in terms of SNV frequency variation through time because it is correlated with MAG coverage). Total number of polymorphic SNVs (excluding fixed single-nucleotide substitutions) is reported in the y-axis title.

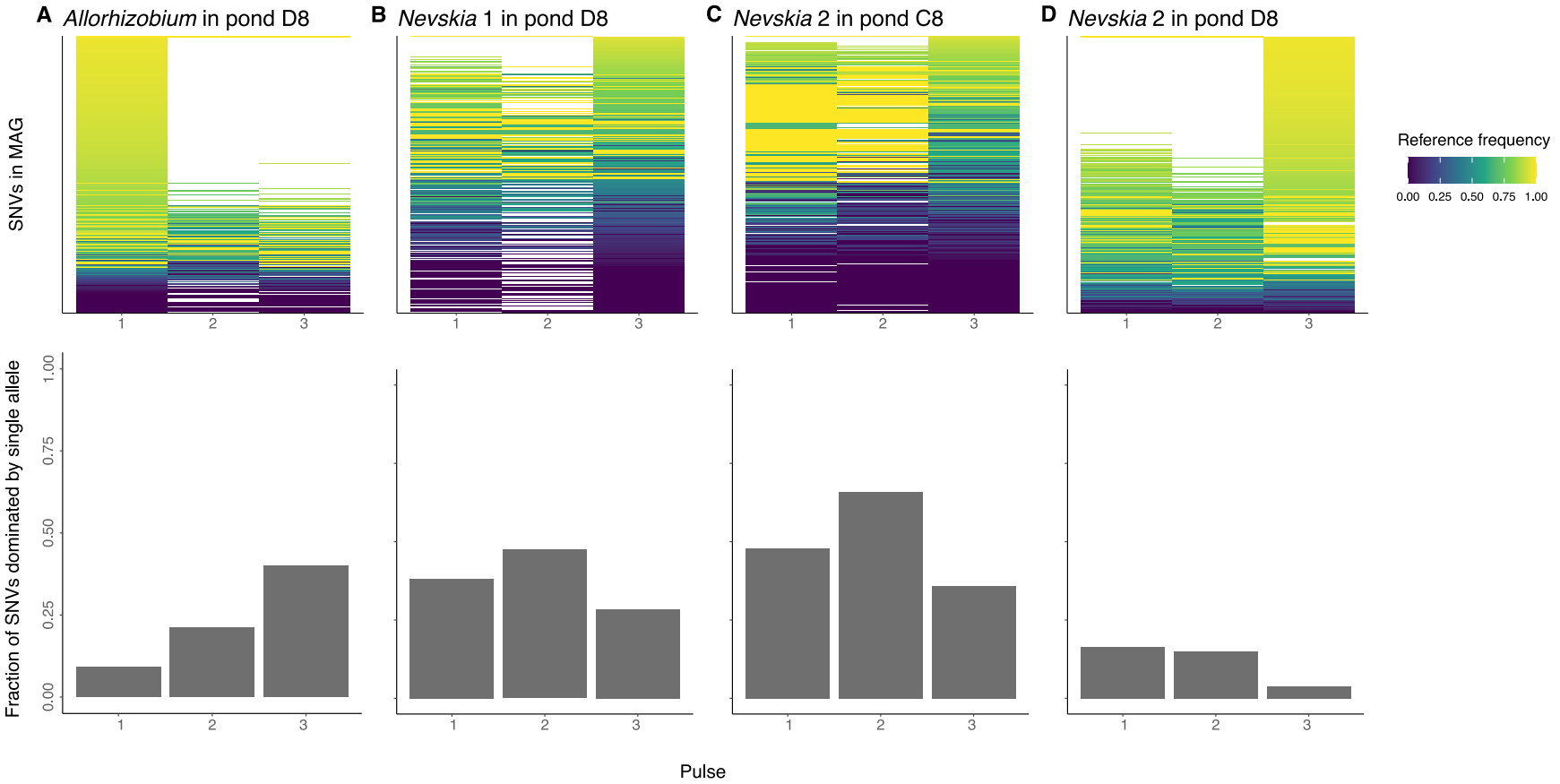


Figure IV.S4 Reference frequency within SNV positions (top) and fraction of SNVs dominated by a single allele (bottom) in populations that exhibited variation in MAG coverage proportional to total SNVs/Mbp and thus prevented assessment of temporal changes after pulses 1, 2 or 3. Only positions with minimum coverage of 5X are shown.

Supplementary tables

Table IV.S1 List of 31 MAGs with PRC score higher than 0.07 or between -0.007 and 0.007, their NCBI accession number, their taxonomic annotation and EPSPS putative classification (according to https://ppuigbo.me/programs/EPSPSClass/).

MAG id	Taxonomy (checkM)	Classification (GTDB-Tk)	EPSPS putative classification	PRC score	SNV profile accessed	Total ARGs in MAG (RGI strict hits)
C4_MAG_00010	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria	d__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Caulobacterales;f__Hyphomonadaceae;g__Aquidulcibacter;s__	sensitive	High (0.0703)	yes	0
C4_MAG_00090	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhizobiales;f__Rhizobiaceae	d__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhizobiales;f__Rhizobiaceae;g__Allorhizobium;s__	resistant	High (0.0998)	yes	1
C8_MAG_00031	k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Xanthomonadales;f__Sinobacteraceae	d__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Nevskiales;f__Nevskiaceae;g__Nevskia;s__	resistant	High (0.1381)	yes	1
C8_MAG_00032	k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Oxalobacteraceae	d__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Burkholderiales;f__Burkholderiaceae;g__SYFN01;s__	unclassified	High (0.1804)	yes	2
C8_MAG_00048	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhodospirillales;f__Rhodospirillaceae;g__Azospirillum	d__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Azospirillales;f__Azospirillaceae;g__Niveispirillum;s__	resistant	High (0.0867)	yes	2
C8_MAG_00052	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhizobiales	d__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Micropepsales;f__Micropepsaceae;g__UBA2784;s__	resistant	High (0.1029)	yes	0
D1_MAG_00019	k__Bacteria;p__Verrucomicrobia;c__Verrucomicrobiae;o__Verrucomicrobiales;f__Verrucomicrobiaceae	d__Bacteria;p__Verrucomicrobiota;c__Verrucomicrobiae;o__Verrucomicrobiales;f__Verrucomicrobiaceae;g__Prosthecobacter;s__	resistant	Low (0.0003)	yes	1
D4_MAG_00049	k__Bacteria;p__Actinobacteria;c__Actinobacteria;o__Actinomycetales;f__Streptomycetaceae	d__Bacteria;p__Actinobacteriota;c__Actinomycetia;o__Nanopelagicales;f__Nanopelagicaceae;g__Planktophila;s__	sensitive	Low (0.0001)	yes	0

D8_MAG_00048	k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Xanthomonadales;f__Sinobacteraceae	d__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Nevskiales;f__Nevskiaceae;g__Nevskia;s__	resistant	High (0.1235)	yes	0
D8_MAG_00058	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Sphingomonadales;f__Sphingomonadaceae;g__Novosphingobium	d__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Sphingomonadales;f__Sphingomonadaceae;g__Novosphingobium;s__	resistant	High (0.3300)	yes	1
E1_MAG_00066	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhodospirillales;f__Acetobacteraceae	d__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Acetobacterales;f__Acetobacteraceae;g__Roseococcus;s__	resistant	High (0.0722)	yes	2
C8_MAG_00007	k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Oxalobacteraceae	d__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Burkholderiales;f__Burkholderiaceae;g__Undibacterium;s__Undibacterium sp014284235	sensitive	High (0.0727)	not for this study	2
C1_MAG_00064	k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Comamonadaceae	d__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Burkholderiales;f__Burkholderiaceae;g__Polaromonas;s__	unclassified	Low (-0.0004)	not enough coverage	0
C4_MAG_00016	k__Bacteria;p__Actinobacteria;c__Actinobacteria;o__Actinomycetales	d__Bacteria;p__Actinobacteriota;c__Actinomycetia;o__Nanopelagicales;f__UBA5976;g__UBA5976;s__	unclassified	Low (-0.0004)	not enough coverage	0
C4_MAG_00039	k__Bacteria;p__Proteobacteria	d__Bacteria;p__Bdellovibrionota_C;c__UBA2361;o__UBA2361;f__UBA2361;g__OMII01;s__	sensitive	Low (0.0007)	not enough coverage	1
C4_MAG_00042	k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Oxalobacteraceae	d__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Burkholderiales;f__Burkholderiaceae;g__SYFN01;s__	sensitive	Low (0.0005)	not enough coverage	1
C4_MAG_00063	k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Rubrivivax	d__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Burkholderiales;f__Burkholderiaceae;g__Rubrivivax;s__	sensitive	Low (0.0005)	not enough coverage	2
C8_MAG_00011	k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Comamonadaceae	d__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Burkholderiales;f__Burkholderiaceae;g__Ramlibacter;s__	sensitive	High (0.1062)	not enough coverage	2

C8_MAG_00015	k__Bacteria;p__Bacteroidetes;c__Flavobacteriia;o__Flavobacteriales	d__Bacteria;p__Bacteroidota;c__Bacteroidia;o__Flavobacteriales;f__Schleiferiaceae;g__TMED14;s__	sensitive	Low (0.0007)	not enough coverage	0
C8_MAG_00039	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhizobiales	d__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhizobiales;f__Beijerinckiaceae;g__;s__	unclassified	High (0.0712)	not enough coverage	1
C8_MAG_00042	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhizobiales	d__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhizobiales;f__Beijerinckiaceae;g__;s__	resistant	High (0.0708)	not enough coverage	1
D1_MAG_00052	k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Rubrivivax	d__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Burkholderiales;f__Burkholderiaceae;g__Rubrivivax;s__	resistant	Low (-0.00002)	not enough coverage	0
D1_MAG_00060	k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Comamonadaceae	d__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Burkholderiales;f__Burkholderiaceae;g__Rhodoferax;s__	unclassified	Low (-0.0004)	not enough coverage	1
D4_MAG_00034	k__Bacteria;p__Verrucomicrobia;c__Opitutae;o__Opitutales;f__Opitutaceae	d__Bacteria;p__Verrucomicrobiota;c__Verrucomicrobiae;o__Opitutales;f__Opitutaceae;g__UBA6669;s__	unclassified	Low (0.0007)	not enough coverage	0
D4_MAG_00063	k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Comamonadaceae	d__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Burkholderiales;f__Burkholderiaceae;g__Limnohabitans_A;s__Limnohabitans_Asp003063375	sensitive	Low (-0.0007)	not enough coverage	0
D4_MAG_00077	k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Comamonadaceae	d__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Burkholderiales;f__Burkholderiaceae;g__CAINMN01;s__	unclassified	Low (0.0003)	not enough coverage	1
D4_MAG_00078	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhizobiales	d__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhizobiales;f__Beijerinckiaceae;g__;s__	resistant	Low (-0.0002)	not enough coverage	0
D8_MAG_00016	k__Bacteria;p__Verrucomicrobia;c__Verrucomicrobiae;o__Verrucomicrobiales;f__Verrucomicrobiaceae	d__Bacteria;p__Verrucomicrobiota;c__Verrucomicrobiae;o__Verrucomicrobiales;f__Verrucomicrobiaceae;g__Prosthecobacter;s__	resistant	High (0.1134)	not enough coverage	3
D8_MAG_00074	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhodobacterales;f__Rhodobacteraceae;g__Rhodobacter	d__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhodobacterales;f__Rhodobacteraceae;g__Tabriziicola;s__	resistant	High (0.1122)	not enough coverage	0

D8_MAG_00085	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Sphingomonadales;f__Erythrobacteraceae	d__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Sphingomonadales;f__Sphingomonadaceae;g__Erythrobacter;s__	resistant	High (0.1390)	not enough coverage	2
E1_MAG_00024	k__Bacteria;p__Bacteroidetes	d__Bacteria;p__Bacteroidota;c__Bacteroidia;o__NS11-12g;f__UKL13-3;g__B1;s__	unclassified	Low (0.0001)	not enough coverage	0

Table IV.S2 SNV and MAG coverage summary of samples selected for SNV profiling

Sample	MAG label	MAG coverage	MAG breadth of coverage	Total SNVs (excluding SNS)	Median SNV position coverage	Median frequency of minor allele	SNVs/Mbp	Position in Figure IV.2	Results summary of intraspecific diversity variation
Prosthecobacter in pond C8 in pulse 1	D1_MAG_00019	5	89%	461	7	0.10	91	A	Steady diversity
Prosthecobacter in pond C8 in pulse 2	D1_MAG_00019	21	89%	113	23	0.00	22	A	Steady diversity
Prosthecobacter in pond C8 in pulse 3	D1_MAG_00019	7	86%	58	11	0.00	11	A	Steady diversity
SYFN01 in pond D8 in pulse 1	C8_MAG_00032	27	85%	78	28	0.00	20	B	Steady diversity
SYFN01 in pond D8 in pulse 2	C8_MAG_00032	65	86%	25	19	0.00	6	B	Steady diversity
SYFN01 in pond D8 in pulse 3	C8_MAG_00032	36	86%	111	33	0.00	28	B	Steady diversity
Roseococcus in pond D8 in pulse 1	E1_MAG_00066	7	95%	106	9	0.15	23	C	Steady diversity
Roseococcus in pond D8 in pulse 2	E1_MAG_00066	8	95%	86	10	0.00	19	C	Steady diversity
Roseococcus in pond D8 in pulse 3	E1_MAG_00066	32	95%	262	36	0.05	57	C	Steady diversity
Novosphingobium in pond D8 in pulse 1	D8_MAG_00058	10	100%	190	11	0.18	44	D	Steady diversity
Novosphingobium in pond D8 in pulse 2	D8_MAG_00058	52	100%	220	53	0.06	51	D	Steady diversity
Novosphingobium in pond D8 in pulse 3	D8_MAG_00058	15	100%	434	17	0.12	101	D	Steady diversity
UBA2784 in pond D8 in pulse 1	C8_MAG_00052	8	93%	3676	11	0.22	1031	E	Steady diversity
UBA2784 in pond D8 in pulse 2	C8_MAG_00052	8	91%	3364	10	0.25	943	E	Steady diversity
UBA2784 in pond D8 in pulse 3	C8_MAG_00052	24	96%	11717	24	0.17	3285	E	Steady diversity

UBA2784 in pond C8 in pulse 1	C8_MAG_00052	2	78%	323	6	0.33	91	F	Steady diversity
UBA2784 in pond C8 in pulse 2	C8_MAG_00052	2	78%	272	6	0.40	76	F	Steady diversity
UBA2784 in pond C8 in pulse 3	C8_MAG_00052	2	77%	645	6	0.33	181	F	Steady diversity
Nevskia 1 in pond C8 in pulse 1	C8_MAG_00031	7	97%	6508	9	0.33	1386	G	Reduced diversity in one timepoint
Nevskia 1 in pond C8 in pulse 2	C8_MAG_00031	13	92%	1478	16	0.13	315	G	Reduced diversity in one timepoint
Nevskia 1 in pond C8 in pulse 3	C8_MAG_00031	68	100%	7768	78	0.08	1654	G	Reduced diversity in one timepoint
Aquidulcibacter in pond D8 in pulse 1	C4_MAG_00010	45	100%	9754	35	0.32	2943	H	Reduced diversity in one timepoint
Aquidulcibacter in pond D8 in pulse 2	C4_MAG_00010	103	100%	10864	82	0.26	3277	H	Reduced diversity in one timepoint
Aquidulcibacter in pond D8 in pulse 3	C4_MAG_00010	6	99%	400	8	0.22	121	H	Reduced diversity in one timepoint
Planktophila in pond D4 in pulse 1	D4_MAG_00049	10	100%	305	10	0.33	247	I	Reduced diversity in one timepoint
Planktophila in pond D4 in pulse 2	D4_MAG_00049	13	100%	305	11	0.29	247	I	Reduced diversity in one timepoint
Planktophila in pond D4 in pulse 3	D4_MAG_00049	18	100%	347	16	0.20	281	I	Reduced diversity in one timepoint
Allorhizobium in pond C8 in pulse 1	C4_MAG_00090	15	91%	921	17	0.07	213	J	Increased diversity in one timepoint
Allorhizobium in pond C8 in pulse 2	C4_MAG_00090	9	90%	473	12	0.00	109	J	Increased diversity in one timepoint
Allorhizobium in pond C8 in pulse 3	C4_MAG_00090	4	84%	1505	6	0.33	348	J	Increased diversity in one timepoint
Niveispirillum in pond C8 in pulse 1	C8_MAG_00048	11	98%	401	12	0.00	77	K	Increased diversity in one timepoint
Niveispirillum in pond C8 in pulse 2	C8_MAG_00048	23	98%	2444	22	0.09	470	K	Increased diversity in one timepoint
Niveispirillum in pond C8 in pulse 3	C8_MAG_00048	17	100%	16237	15	0.25	3126	K	Increased diversity in one timepoint
Allorhizobium in pond D8 in pulse 1	C4_MAG_00090	97	92%	12959	77	0.08	2999	NA	Variation in MAG coverage

Allorhizobium in pond D8 in pulse 2	C4_MAG_00090	3	77%	1212	6	0.38	280	NA	Variation in MAG coverage
Allorhizobium in pond D8 in pulse 3	C4_MAG_00090	4	85%	721	7	0.30	167	NA	Variation in MAG coverage
Nevskia 1 in pond D8 in pulse 1	C8_MAG_00031	2	56%	1738	7	0.29	370	NA	Variation in MAG coverage
Nevskia 1 in pond D8 in pulse 2	C8_MAG_00031	1	47%	684	6	0.33	146	NA	Variation in MAG coverage
Nevskia 1 in pond D8 in pulse 3	C8_MAG_00031	25	83%	10186	22	0.19	2169	NA	Variation in MAG coverage
Nevskia 2 in pond C8 in pulse 1	D8_MAG_00048	5	75%	3544	9	0.25	762	NA	Variation in MAG coverage
Nevskia 2 in pond C8 in pulse 2	D8_MAG_00048	3	62%	1307	8	0.17	281	NA	Variation in MAG coverage
Nevskia 2 in pond C8 in pulse 3	D8_MAG_00048	21	82%	8427	19	0.18	1813	NA	Variation in MAG coverage
Nevskia 2 in pond D8 in pulse 1	D8_MAG_00048	7	85%	2631	10	0.22	566	NA	Variation in MAG coverage
Nevskia 2 in pond D8 in pulse 2	D8_MAG_00048	4	82%	1416	7	0.33	305	NA	Variation in MAG coverage
Nevskia 2 in pond D8 in pulse 3	D8_MAG_00048	76	99%	9027	93	0.07	1942	NA	Variation in MAG coverage

CHAPTER V : CONCLUSIONS AND FUTURE PERSPECTIVES

As highlighted in chapter I, although pollution is a major driver of biodiversity loss whose projected environmental impacts make it a critical policy target, this issue has received low research attention relative to its importance when compared to other global biodiversity threats, such as climate change (Mazor et al., 2018). The studies presented in this thesis contribute to revealing the complexity of bacterial responses to agrochemical contamination while integrating ecology and evolution with ecotoxicology, claimed to be important in the assessment and prediction of contaminants' impacts on populations, communities and ecosystems (Gessner & Tlili, 2016; M. Oziolor, De Schamphelaere, & Matson, 2016).

Chapters II and III contributed to exploring community responses while focusing on species composition, functions and genes, while chapter IV contributed to exploring populational responses to the most widely used herbicide on the planet. Consequences of community changes to ecosystem functions were explored in chapter II through the investigation of usage of carbon sources by the microbial community. Moreover, upcoming avenues of this work will provide an overview of gene fluctuations and mutations through time that may indicate a potential alteration in ecosystem functions provided by bacterial species. Although the investigation of genes predicted from metagenomic data does not correspond to actually expressed functions it may indicate potential changes in ecosystem functions.

By focusing on the response of communities originally from a natural environment, the results of chapter II bring an additional perspective to traditional ecotoxicology studies that focus on single species, as community properties may affect how they react to a disturbance (Shade et al., 2012). Furthermore, chapter II showed how the observation of community resilience may vary whether the composition is described in terms of ASVs or any other broader-scale phylogenetic grouping. By showing that the originally existing ASVs never returned to be part of the community, but were replaced by close relatives, this study demonstrates how results may be variable according to the taxonomic scale of bacterial biodiversity studies. Research on microbial diversity has been revolutionized by the use of next-generation sequencing technologies (D. P. Smith & Peay, 2014) and, more recently, an alternative method for detecting diversity units based on ASVs has

been suggested to replace operational taxonomic units (OTUs) to perform reproducible analyses at the finest possible taxonomic level (Benjamin J Callahan et al., 2017). A clear limitation of the *de novo* OTUs prediction method is its absence of reproducibility: as it is based on a sequence similarity threshold (usually 97%) of a cluster of reads, OTUs inferred from different datasets cannot be compared, which differs from ASVs whose prediction is also independent of reference databases and that are comparable across studies (Benjamin J Callahan et al., 2017). Nevertheless, the biological relevance of ASVs fine-resolution has also been debated, as the method may increase the risk of splitting a single bacterial genome into separate clusters (Schloss, 2021). As ASV inference methods become prevalent but there is not yet a consensus regarding their biological significance, the approach used in chapter 2 (i.e. reanalyzing ASV data after clustering sequences according to phylogenetic distances) provides a workaround to assess the differences between more and less refined taxonomy levels.

While chapter II provided an overview of community responses to a neonicotinoid insecticide and a GBH, using a combination of methods such as flow cytometry, community-level physiological profiling and 16S rRNA sequencing, the next two chapters focused on the GBH treatments, which produced significant ecological changes in the bacterioplankton, through metagenomic sequencing to explore changes in the composition of genes within communities (chapter III) and allele fluctuations within populations (chapter IV). The metagenomic approach characterizes environmental samples beyond the composition and diversity of species, as it allows the prediction of genes as well as the reconstruction of genomes and inference of intraspecific diversity, being thus useful for genetic and evolutionary studies.

The use of metagenomic sequencing allowed the non-targeted prediction of genes and the estimation of ARGs frequency in the experimental communities in chapter III. The study provided the first evidence of the cross-selection of ARGs by a GBH in an aquatic environment and highlighted that efflux pumps explained bacterial abundance in the presence of high concentrations of the herbicide better than the resistance mutation in the glyphosate-target enzyme. To verify the risk of ARGs selection in an environmental context, future studies could perform a comparative analysis of natural aquatic systems with different degrees of contamination, where concentrations of residual glyphosate tend

to be lower and more variable. An example from soil microbiome research is a field study done across 21 agricultural lands in China with a different history of glyphosate exposure, and the authors confirmed the correlation between the relative abundance of ARGs, quantified through qPCR, and the concentration of glyphosate residues in soil (Liao et al., 2021).

In chapter IV, a few genomes assembled from metagenomes (i.e. MAGs) with similar ecological responses to GBH contamination through time were studied regarding their evolutionary dynamics. Differently than expected, allele variation analyses showed that intraspecific diversity varied idiosyncratically among these MAGs, showing that ecological changes within short timescales are not always accompanied by directional evolutionary changes. Furthermore, this chapter provided one example of a genome-wide selective sweep, supporting the evolutionary theory proposed in the ecotype model, according to which bacterial differentiation arises from occasional diversity purge as a consequence of selective pressure (Cohan, 2001; Cohan & Perry, 2007). Other studies have already provided support for the ecotype model, such as a 9-year metagenomic research in a freshwater lake that observed a genome-wide sweep in 1 out of 30 reconstructed genomes (Bendall et al., 2016) and the comparative study of 90k genomes that revealed a discontinuous genetic diversity among them (Jain et al., 2018). Differently from these two studies, chapter IV contributed as an experimental study where a selective pressure is known and thus, decreasing the probability that the sweep happened due to drift or random demographic changes.

This thesis evidences how ecology, genetics and evolution explain interconnected processes driving microbial responses against environmental changes. Chapters II, III and IV provide novel contributions to the scientific community on how freshwater bacterial communities respond to short-term agrochemical contamination while using a large-scale experimental approach and high-throughput sequencing.

As a continuation of this thesis, we performed a reproduction of the experiment in the LEAP platform in 2020 and 2021 to verify if our core findings are replicable. These experiments will complement chapter III by exploring how GBH drives phenotypic and genotypic changes regarding antimicrobial resistance in bacterial strains and chapter IV by investigating evolutionary responses of glyphosate sensitive species in a replicated

system. In combination, the studies here presented and ongoing investigations derived from them are contributing to the advancement of ecotoxicological studies as they highlight the complexity of microbial responses against a contaminant.

References

Alcock, B. P., Raphenya, A. R., Lau, T. T. Y., Tsang, K. K., Bouchard, M., Edalatmand, A., … McArthur, A. G. (2020). CARD 2020: antibiotic resistome surveillance with the comprehensive antibiotic resistance database. *Nucleic Acids Research*, *48*(D1), D517–D525. https://doi.org/10.1093/nar/gkz935

Alexander, A. C., Luiker, E., Finley, M., & Culp, J. M. (2016). Mesocosm and Field Toxicity Testing in the Marine Context. In *Marine Ecotoxicology* (pp. 239–256). Elsevier Inc. https://doi.org/10.1016/b978-0-12-803371-5.00008-4

Alexander, A. C., Luis, A. T., Culp, J. M., Baird, D. J., & Cessna, A. J. (2013). Can nutrients mask community responses to insecticide mixtures? *Ecotoxicology*, *22*(7), 1085–1100. https://doi.org/10.1007/s10646-013-1096-3

Allison, S. D., & Martiny, J. B. H. (2008). Resistance, resilience, and redundancy in microbial communities. *Proceedings of the National Academy of Sciences of the USA*, *105*, 11512–11519. https://doi.org/10.1073/pnas.0801925105

Alneberg, J., Bjarnason, B. S., de Bruijn, I., Schirmer, M., Quick, J., Ijaz, U. Z., … Quince, C. (2014). Binning metagenomic contigs by coverage and composition. *Nature Methods*, *11*(11), 1144–1146. https://doi.org/10.1038/nmeth.3103

Altenburger, R., Backhaus, T., Boedeker, W., Faust, M., & Scholze, M. (2013). Simplifying complexity: Mixture toxicity assessment in the last 20 years. *Environmental Toxicology and Chemistry*, *32*(8), 1685–1687. https://doi.org/10.1002/etc.2294

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, *215*(3), 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2

Anderson, M. J. (2001). A new method for non-parametric multivariate analysis of variance, *26*, 32–46.

Anderson, M. J. (2006). Distance-based tests for homogeneity of multivariate dispersions. *Biometrics*, *62*(1), 245–253. https://doi.org/10.1111/j.1541-0420.2005.00440.x

Auber, A., Travers-Trolet, M., Villanueva, M.-C., & Ernande, B. (2017). A new application of principal response curves for summarizing abrupt and cyclic shifts of

communities over space. *Ecosphere*, *8*(12), e02023. https://doi.org/10.1002/ecs2.2023

Baker-Austin, C., Wright, M. S., Stepanauskas, R., & McArthur, J. V. (2006). Co-selection of antibiotic and metal resistance. *Trends in Microbiology*, *14*(4), 176–182. https://doi.org/10.1016/j.tim.2006.02.006

Baker, L. F., Mudge, J. F., Thompson, D. G., Houlahan, J. E., & Kidd, K. A. (2016). The combined influence of two agricultural contaminants on natural communities of phytoplankton and zooplankton. *Ecotoxicology*, *25*(5), 1021–1032. https://doi.org/10.1007/s10646-016-1659-1

Barbosa da Costa, N., Fugère, V., Hébert, M., Xu, C. C. Y., Barrett, R. D. H., Beisner, B. E., … Shapiro, B. J. (2021). Resistance, resilience, and functional redundancy of freshwater bacterioplankton communities facing a gradient of agricultural stressors in a mesocosm experiment. *Molecular Ecology*, *30*(19), 4771–4788. https://doi.org/10.1111/mec.16100

Barbosa da Costa, N., Hébert, M.-P., Fugère, V., Terrat, Y., Fussmann, G. F., Gonzalez, A., & Shapiro, B. J. (2022). A Glyphosate-Based Herbicide Cross-Selects for Antibiotic Resistance Genes in Bacterioplankton Communities. *MSystems*, *7*(2), e01482-21. https://doi.org/10.1128/msystems.01482-21

Beaulieu, J. J., DelSontro, T., & Downing, J. A. (2019). Eutrophication will increase methane emissions from lakes and impoundments during the 21st century. *Nature Communications*, *10*(1), 3–7. https://doi.org/10.1038/s41467-019-09100-5

Beauséjour, R., Handa, I. T., Lechowicz, M. J., Gilbert, B., & Vellend, M. (2015). Historical anthropogenic disturbances influence patterns of non-native earthworm and plant invasions in a temperate primary forest. *Biological Invasions*, *17*(4), 1267–1281. https://doi.org/10.1007/s10530-014-0794-y

Benbrook, C. M. (2016). Trends in glyphosate herbicide use in the United States and globally. *Environmental Sciences Europe*, *28*(1), 1–15. https://doi.org/10.1186/s12302-016-0070-0

Bendall, M. L., Stevens, S. L., Chan, L.-K., Malfatti, S., Schwientek, P., Tremblay, J., … Malmstrom, R. R. (2016). Genome-wide selective sweeps and gene-specific sweeps in natural bacterial populations. *The ISME Journal*, *10*(7), 1–13.

https://doi.org/10.1038/ismej.2015.241

Benton, T. G., Solan, M., Travis, J. M. J., & Sait, S. M. (2007). Microcosm experiments can inform global ecological problems. *Trends in Ecology and Evolution*, *22*(10), 516–521. https://doi.org/10.1016/j.tree.2007.08.003

Berggren, M., Lapierre, J. F., & del Giorgio, P. A. (2012). Magnitude and regulation of bacterioplankton respiratory quotient across freshwater environmental gradients. *The ISME Journal*, *6*(5), 984–993. https://doi.org/10.1038/ismej.2011.157

Berman, M. C., Llames, M. E., Minotti, P., Fermani, P., Quiroga, M. V., Ferraro, M. A., … Zagarese, H. E. (2020). Field evidence supports former experimental claims on the stimulatory effect of glyphosate on picocyanobacteria communities. *Science of the Total Environment*, *701*, 134601. https://doi.org/10.1016/j.scitotenv.2019.134601

Birk, S., Chapman, D., Carvalho, L., Spears, B. M., Andersen, H. E., Argillier, C., … Hering, D. (2020). Impacts of multiple stressors on freshwater biota across spatial scales and ecosystems. *Nature Ecology & Evolution*, *4*(8), 1060–1068. https://doi.org/10.1038/s41559-020-1216-4

Bloomfield, J. P., Williams, R. J., Gooddy, D. C., Cape, J. N., & Guha, P. (2006). Impacts of climate change on the fate and behaviour of pesticides in surface and groundwater-a UK perspective. *Science of the Total Environment*, *369*(1–3), 163–177. https://doi.org/10.1016/j.scitotenv.2006.05.019

Bobay, L., & Ochman, H. (2017). Biological Species Are Universal across Life ' s Domains, *9*(3), 491–501. https://doi.org/10.1093/gbe/evx026

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Brook, D., & Beaton, T. (2015). *Pesticide Contamination of Farm Water Sources*. *Factsheet Ontario Ministry of Agriculture, Food and Rural Affairs*.

Brovini, E. M., Cardoso, S. J., Quadra, G. R., Vilas-Boas, J. A., Paranaíba, J. R., Pereira, R. de O., & Mendonça, R. F. (2021). Glyphosate concentrations in global freshwaters: are aquatic organisms at risk? *Environmental Science and Pollution Research*, (Table 1). https://doi.org/10.1007/s11356-021-14609-8

Callahan, B., McMurdie, P., Rosen, M., Han, A. W., Johnson, A. J. A., & Holmes, S. P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. *Nature Methods*, *13*(7), 581–583. https://doi.org/10.1038/nmeth.3869

Callahan, Ben J., Sankaran, K., Fukuyama, J. A., McMurdie, P. J., & Holmes, S. P. (2016). Bioconductor workflow for microbiome data analysis: From raw reads to community analyses. *F1000Research*, *5*(1492), 1–49. https://doi.org/10.12688/f1000research.8986.2

Callahan, Benjamin J, McMurdie, P. J., & Holmes, S. P. (2017). Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *The ISME Journal*, 1–5. https://doi.org/10.1038/ismej.2017.119

Campbell, B. M., Beare, D. J., Bennett, E. M., Hall-Spencer, J. M., Ingram, J. S. I., Jaramillo, F., … Shindell, D. (2017). Agriculture production as a major driver of the earth system exceeding planetary boundaries. *Ecology and Society*, *22*(4). https://doi.org/10.5751/ES-09595-220408

Carpenter, S. (1996). Microcosm Experiments have Limited Relevance for Community and Ecosystem Ecology. *Ecology*, *77*(3), 677–680.

Carpenter, S. R., Caraco, N. F., Correll, D. L., Howarth, R. W., Sharpley, A. N., & Smith, V. H. (1998). Nonpoint pollution of surface waters with phosphorus and nitrogen. *Ecological Application*, *8*(January 1998), 559–568. https://doi.org/10.1890/1051-0761(1998)008[0559:NPOSWW]2.0.CO;2

Carpenter, S. R., Stanley, E. H., & Vander Zanden, M. J. (2011). State of the world's freshwater ecosystems: Physical, chemical, and biological changes. *Annual Review of Environment and Resources*, *36*, 75–99. https://doi.org/10.1146/annurev-environ-021810-094524

Caruso, V., Song, X., Asquith, M., & Karstens, L. (2019). Performance of Microbiome Sequence Inference Methods in Environments with Varying Biomass. *MSystems*, *4*(1), 1–19. https://doi.org/10.1128/msystems.00163-18

CCME. (2007). *Canadian Water Quality Guidelines for Protection of Aquatic Life: Imidacloprid*. *Scientific Supporting Document*. Winnipeg, Manitoba.

CCME. (2012). *Canadian Water Quality Guidelines for the Protection of Aquatic Life: Glyphosate*. *Scientific Supporting Document*. Winnipeg, Manitoba.

Cedergreen, N., & Rasmussen, J. J. (2017). Low dose effects of pesticides in the aquatic environment. In *ACS Symposium Series* (Vol. 1249, pp. 167–187). https://doi.org/10.1021/bk-2017-1249.ch012

Chará-Serna, A. M., Epele, L. B., Morrissey, C. A., & Richardson, J. S. (2019). Nutrients and sediment modify the impacts of a neonicotinoid insecticide on freshwater community structure and ecosystem functioning. *Science of the Total Environment*, *692*, 1291–1303. https://doi.org/10.1016/j.scitotenv.2019.06.301

Chaumeil, P. A., Mussig, A. J., Hugenholtz, P., & Parks, D. H. (2020). GTDB-Tk: A toolkit to classify genomes with the genome taxonomy database. *Bioinformatics*, *36*(6), 1925–1927. https://doi.org/10.1093/bioinformatics/btz848

Cheng, F. Y., & Basu, N. B. (2017). Biogeochemical Hotspots: Role of Small Water Bodies in Landscape Nutrient Processing, 1–12.

Clements, W. H., & Rohr, J. R. (2009). Community responses to contaminants: using basic ecological principles to predict ecotoxicological effects. *Environmental Toxicology and Chemistry*, *28*(9), 1789–1800. https://doi.org/10.1897/09-140.1

Cohan, F. M. (2001). Bacterial species and speciation. *Systems Biology*, *50*(4), 513–524.

Cohan, F. M. (2016). Bacterial speciation: Genetic sweeps in bacterial species. *Current Biology*, *26*(3), R112–R115. https://doi.org/10.1016/j.cub.2015.10.022

Cohan, F. M. (2019). Systematics: The Cohesive Nature of Bacterial Species Taxa. *Current Biology*, *29*(5), R169–R172. https://doi.org/10.1016/j.cub.2019.01.033

Cohan, F. M., & Perry, E. B. (2007). A Systematics for Discovering the Fundamental Units of Bacterial Diversity. *Current Biology*, *17*(10), 373–386. https://doi.org/10.1016/j.cub.2007.03.032

Cole, J. J., Prairie, Y. T., Caraco, N. F., Mcdowell, W. H., Tranvik, L. J., Striegl, R. G., … Melack, J. (2007). Plumbing the Global Carbon Cycle : Integrating Inland Waters into the Terrestrial Carbon Budget, 171–184. https://doi.org/10.1007/s10021-006-9013-8

Coleman, M. L., & Chisholm, S. W. (2010). Ecosystem-specific selection pressures revealed through comparative population genomics. *Proc. Natl. Acad. Sci. Unit. States Am.*, *107*(43), 18634–18639. https://doi.org/10.1073/pnas.1009480107/-

/DCSupplemental.www.pnas.org/cgi/doi/10.1073/pnas.1009480107

Cordero, O. X., & Polz, M. F. (2014). Explaining microbial genomic diversity in light of evolutionary ecology. *Nature Publishing Group*, *12*(4), 263–273. https://doi.org/10.1038/nrmicro3218

Davidson, A. L., & Chen, J. (2004). ATP-binding cassette transporters in bacteria. *Annual Review of Biochemistry*, *73*, 241–268. https://doi.org/10.1146/annurev.biochem.73.011303.073626

Delcher, A. L., Phillippy, A., Carlton, J., & Salzberg, S. L. (2002). Fast algorithms for large-scale genome alignment and comparison, *30*(11), 2478–2483.

Delmont, T. O., Quince, C., Shaiber, A., Esen, Ö. C., Lee, S. T., Rappé, M. S., … Eren, A. M. (2018). Nitrogen-fixing populations of Planctomycetes and Proteobacteria are abundant in surface ocean metagenomes. *Nature Microbiology*, *3*(8), 963. https://doi.org/10.1038/s41564-018-0209-4

DeLorenzo, M. E., Scott, G. I., & Ross, P. E. (2001). Toxicity of pesticides to aquatic microorganisms: A review. *Environmental Toxicology and Chemistry*, *20*(1), 84–98. https://doi.org/10.1002/etc.5620200108

DeSantis, T. Z., Hugenholtz, P., Larsen, N., Rojas, M., Brodie, E. L., Keller, K., … Andersen, G. L. (2006). Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Applied and Environmental Microbiology*, *72*(7), 5069–5072. https://doi.org/10.1128/AEM.03006-05

Doug Hyatt, Gwo-Liang Chen, Philip F LoCascio, Miriam L Land, Frank W Larimer, L. J. H. (2010). Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics*, *11*(119), 1–8. Retrieved from http://dx.doi.org/10.1016/B978-0-12-407863-5.00023-X%5Cnhttp://www.nature.com/doifinder/10.1038/ismej.2009.79%5Cnhttp://www.nature.com/doifinder/10.1038/nature09916%5Cnhttp://dx.doi.org/10.1038/srep25982%5Cnhttp://dx.doi.org/10.1038/ismej.2010.144%5Cnhttp

Duke, S. O., & Powles, S. B. (2008). Glyphosate: a once-in-a-century herbicide. *Pest Management Science*, *63*(11), 1100–1106. https://doi.org/10.1002/ps

Eddy, S. R. (2011). Accelerated profile HMM searches. *PLoS Computational Biology*, *7*(10). https://doi.org/10.1371/journal.pcbi.1002195

EFSA. (2014). *Conclusion on the peer review of the pesticide risk assessment for aquatic organisms for the active substance imidacloprid*. *EFSA Journal* (Vol. 12). https://doi.org/10.2903/j.efsa.2013.3066

EFSA. (2016). *Conclusion on the peer review of the pesticide risk assessment of the active substance glyphosate*. *EFSA Journal* (Vol. 13). https://doi.org/10.2903/j.efsa.2015.4302

Ellis, E. C., Kaplan, J. O., Fuller, D. Q., Vavrus, S., Goldewijk, K. K., & Verburg, P. H. (2013). Used planet: A global history. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(20), 7978–7985. https://doi.org/10.1073/pnas.1217241110

EPA. (2019). Aquatic Life Benchmarks and Ecological Risk Assessments for Registered Pesticides. Retrieved from https://www.epa.gov/pesticide-science-and-assessing-pesticide-risks/aquatic-life-benchmarks-and-ecological-risk#ref_4

Eren, A. M., Esen, Ö. C., Quince, C., Vineis, J. H., Morrison, H. G., Sogin, M. L., & Delmont, T. O. (2015). Anvi'o: an advanced analysis and visualization platform for 'omics data. *PeerJ*, *3*, e1319. https://doi.org/10.7717/peerj.1319

Evans, P. N., Boyd, J. A., Leu, A. O., Woodcroft, B. J., Parks, D. H., Hugenholtz, P., & Tyson, G. W. (2019). An evolving view of methane metabolism in the Archaea. *Nature Reviews Microbiology*, *17*(4), 219–232. https://doi.org/10.1038/s41579-018-0136-7

Falkowski, P. G., Fenchel, T., & Delong, E. F. (2008). The microbial engines that drive earth's biogeochemical cycles. *Science*, *320*(5879), 1034–1039. https://doi.org/10.1126/science.1153213

FAO. (2017). *The future of food and agriculture: trends and challenges*. *The future of food and agriculture: trends and challenges* (Vol. 4). Retrieved from www.fao.org/publications%0Ahttp://www.fao.org/3/a-i6583e.pdf%0Ahttp://siteresources.worldbank.org/INTARD/825826-1111044795683/20424536/Ag_ed_Africa.pdf%0Awww.fao.org/cfs%0Ahttp://www.jstor.org/stable/4356839%0Ahttps://ediss.uni-goettingen.de/bitstream/han

FAO. (2021a). Database collection of the Food and Agriculture Organization of the United Nations, inputs/Nutrient nitrogen and phosphate. Retrieved from

https://www.fao.org/faostat/en/data/

FAO. (2021b). Database collection of the Food and Agriculture Organization of the United Nations, inputs/pesticides use.

Fernandes, G., Carolina, V., Camotti, M., Gerónimo, E. De, Labanowski, J., Damian, O., … Rheinheimer, D. (2019). Indiscriminate use of glyphosate impregnates river epilithic biofilms in southern Brazil. *Science of the Total Environment*, *651*, 1377–1387. https://doi.org/10.1016/j.scitotenv.2018.09.292

Flood, S. L., & Burkholder, J. A. M. (2018). Imbalanced nutrient regimes increase Prymnesium parvum resilience to herbicide exposure. *Harmful Algae*, *75*, 57–74. https://doi.org/10.1016/j.hal.2018.04.006

Foley, J. A., DeFries, R., Asner, G. P., Barford, C., Bonan, G., Carpenter, S. R., … Snyder, P. K. (2005). Global consequences of land use. *Science*, *309*(5734), 570–574. https://doi.org/10.1126/science.1111772

Foley, J. A., Ramankutty, N., Brauman, K. A., Cassidy, E. S., Gerber, J. S., Johnston, M., … Zaks, D. P. M. (2011). Solutions for a cultivated planet. *Nature*, *478*(7369), 337–342. https://doi.org/10.1038/nature10452

Fraser, C., Alm, E. J., Polz, M. F., Spratt, B. G., & Hanage, W. P. (2009). The Bacterial Species Challenge : Ecological Diversity. *Science*, *323*(February), 741–746. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/19197054

Fugère, V., Hébert, M., Barbosa da Costa, N., Xu, C. C. Y., Barrett, R. D. H., Beisner, B. E., … Gonzalez, A. (2020). Community rescue in experimental phytoplankton communities facing severe herbicide pollution. *Nature Ecology & Evolution*, *4*, 578–588. https://doi.org/10.1038/s41559-020-1134-5

Funke, T., Han, H., Healy-Fried, M. L., Fischer, M., & Schönbrunn, E. (2006). Molecular basis for the herbicide resistance of Roundup Ready crops. *Proceedings of the National Academy of Sciences of the USA*, *103*(35), 13010–13015. https://doi.org/10.1073/pnas.0603638103

Fussmann, G. F., Loreau, M., & Abrams, P. A. (2007). Eco-evolutionary dynamics of communities and ecosystems. *Functional Ecology*, *21*(3), 465–477. https://doi.org/10.1111/j.1365-2435.2007.01275.x

Garland, J. L., Mills, A. L., & Young, J. S. (2001). Relative effectiveness of kinetic

analysis vs single point readings for classifying environmental samples based on community-level physiological profiles (CLPP). *Soil Biology and Biochemistry*, *33*(7–8), 1059–1066. https://doi.org/10.1016/S0038-0717(01)00011-6

Gasol, J. M., & Del Giorgio, P. A. (2000). Using flow cytometry for counting natural planktonic bacteria and understanding the structure of planktonic bacterial communities. *Scientia Marina*, *64*(2), 197–224. https://doi.org/10.3989/scimar.2000.64n2197

Gessner, M. O., & Tlili, A. (2016). Fostering integration of freshwater ecology with ecotoxicology. *Freshwater Biology*, *61*(12), 1991–2001. https://doi.org/10.1111/fwb.12852

Geyer, R. L., Smith, G. R., & Rettig, J. E. (2016). Effects of Roundup formulations, nutrient addition, and Western mosquitofish (Gambusia affinis) on aquatic communities. *Environmental Science and Pollution Research*, *23*(12), 11729–11739. https://doi.org/10.1007/s11356-016-6381-2

Girvan, M., Campbell, C., Killham, K., Prosser, J., & Glover, L. (2005). Bacterial diversity promotes community stability and functional resilience after perturbation. *Environmental Microbiology*, *7*(3), 301–313. https://doi.org/10.1111/j.1462-2920.2004.00695.x

Gullberg, E., Cao, S., Berg, O. G., Ilbäck, C., Sandegren, L., Hughes, D., & Andersson, D. I. (2011). Selection of resistant bacteria at very low antibiotic concentrations. *PLoS Pathogens*, *7*(7), 1–9. https://doi.org/10.1371/journal.ppat.1002158

Gurdev S. Khush. (2001). Green revolution: the way forward. *Nature Reviews Genetics*, *2*(October), 815–822.

Healy-Fried, M. L., Funke, T., Priestman, M. A., Han, H., & Scho, E. (2007). Structural Basis of Glyphosate Tolerance Resulting from Mutations of Pro 101 in Escherichia coli. *The Journal of Biological Chemistry*, *282*(45), 32949–32955. https://doi.org/10.1074/jbc.M705624200

Hébert, M.-P., Fugère, V., Beisner, B. E., Barbosa da Costa, N., Barrett, R. D. H., Bell, G., … Fussmann, G. F. (2021). Widespread agrochemicals differentially affect zooplankton biomass and community structure. *Ecological Applications*. https://doi.org/10.1002/eap.2423

Hébert, M.-P., Fugère, V., & Gonzalez, A. (2018). The overlooked impact of rising glyphosate use on phosphorus loading in agricultural watersheds. *Frontiers in Ecology and the Environment*, *3*, 1–9. https://doi.org/10.1002/fee.1985

Hébert, M. P., Fugère, V., Beisner, B. E., Barbosa da Costa, N., Barrett, R. D. H., Bell, G., … Fussmann, G. F. (2021). Widespread agrochemicals differentially affect zooplankton biomass and community structure. *Ecological Applications*, *0*(0), 1–19. https://doi.org/10.1002/eap.2423

Hénault-Ethier, L., Lucotte, M., Moingt, M., Paquet, S., Maccario, S., Smedbol, É., … Labrecque, M. (2017). Herbaceous or Salix miyabeana 'SX64' narrow buffer strips as a means to minimize glyphosate and aminomethylphosphonic acid leaching from row crop fields. *Science of the Total Environment*, *598*, 1177–1186. https://doi.org/10.1016/j.scitotenv.2017.04.104

Hove-Jensen, B., Zechel, D. L., & Jochimsen, B. (2014). Utilization of glyphosate as phosphate source: biochemistry and genetics of bacterial carbon-phosphorus lyase. *Microbiology and Molecular Biology Reviews : MMBR*, *78*(1), 176–197. https://doi.org/10.1128/MMBR.00040-13

Hugerth, L. W., & Andersson, A. F. (2017). Analysing microbial community composition through amplicon sequencing: From sampling to hypothesis testing. *Frontiers in Microbiology*, *8*(SEP), 1–22. https://doi.org/10.3389/fmicb.2017.01561

Huot, Y., Brown, C. A., Potvin, G., Antoniades, D., Baulch, H. M., Beisner, B. E., … Walsh, D. A. (2019). The NSERC Canadian Lake Pulse Network: A national assessment of lake health providing science for water management in a changing climate. *Science of The Total Environment*, *695*, 133668. https://doi.org/10.1016/j.scitotenv.2019.133668

Hutchins, D. A., Walworth, N. G., Webb, E. A., Saito, M. A., Moran, D., McIlvin, M. R., … Fu, F. X. (2015). Irreversibly increased nitrogen fixation in Trichodesmium experimentally adapted to elevated carbon dioxide. *Nature Communications*, *6*. https://doi.org/10.1038/ncomms9155

IPCC. (2022). *Climate Change 2022: Impacts, Adaptation and Vulnerability Working Group II Contribution to the IPCC Sixth Assessment Report*. (H. Pörtner, D. Roberts, M. Tignor, E. Poloczanska, K. Mintenbeck, A. Alegría, … B. Rama, Eds.).

https://doi.org/10.1017/9781009325844.Front

Jain, C., Rodriguez-R, L. M., Phillippy, A. M., Konstantinidis, K. T., & Aluru, S. (2018). High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nature Communications*, *9*(1), 1–8. https://doi.org/10.1038/s41467-018-07641-9

Janßen, R., Skeff, W., Werner, J., Wirth, M. A., Kreikemeyer, B., Schulz-Bull, D., & Labrenz, M. (2019). A Glyphosate Pulse to Brackish Long-Term Microcosms Has a Greater Impact on the Microbial Diversity and Abundance of Planktonic Than of Biofilm Assemblages. *Frontiers in Marine Science*, *6*(December), 1–17. https://doi.org/10.3389/fmars.2019.00758

Jeppesen, E., Kronvang, B., Meerhoff, M., Søndergaard, M., Hansen, K. M., Andersen, H. E., … Olesen, J. E. (2009). Climate Change Effects on Runoff, Catchment Phosphorus Loading and Lake Ecological State, and Potential Adaptations. *Journal of Environment Quality*, *38*(5), 1930. https://doi.org/10.2134/jeq2008.0113

Jeschke, P., & Nauen, R. (2008). Neonicotinoids – from zero to hero in insecticide chemistry. *Pest Management Science*, *64*, 1100–1106. https://doi.org/10.1002/ps.1631

Jost, L. (2006). Entropy and diversity. *Oikos*, *113*(2), 363–375.

Keatley, B. E., Bennett, E. M., Macdonald, G. K., Taranu, Z. E., & Gregory-Eaves, I. (2011). Land-Use Legacies Are Important Determinants of Lake Eutrophication in the Anthropocene. *PLOS One*, *6*(1), e15913. https://doi.org/10.1371/journal.pone.0015913

Khadra, M., Planas, D., Girard, C., & Amyot, M. (2018). Age matters: Submersion period shapes community composition of lake biofilms under glyphosate stress. *Facets*, *3*(1), 934–951. https://doi.org/10.1139/facets-2018-0019

Kim, D., Song, L., Breitwieser, F. P., & Salzberg, S. L. (2016). Centrifuge: rapid and sensitive classification of metagenomic\nsequences. *Genome Res.*, gr.210641.116. https://doi.org/10.1101/gr.210641.116

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

Kraemer, S. A., Barbosa da Costa, N., Shapiro, B. J., Fradette, M., Huot, Y., & Walsh,

D. A. (2020). A large-scale assessment of lakes reveals a pervasive signal of land use on bacterial communities. *The ISME Journal*. https://doi.org/10.1038/s41396-020-0733-0

Kryazhimskiy, S., & Plotkin, J. B. (2008). The population genetics of dN/dS. *PLoS Genetics*, *4*(12). https://doi.org/10.1371/journal.pgen.1000304

Kurenbach, B., Gibson, P. S., Hill, A. M., Bitzer, A. S., Silby, M. W., Godsoe, W., & Heinemann, J. A. (2017). Herbicide ingredients change Salmonella enterica sv. Typhimurium and Escherichia coli antibiotic responses. *Microbiology*, *163*(12), 1791–1801. https://doi.org/10.1099/mic.0.000573

Kurenbach, B., Hill, A. M., Godsoe, W., Van Hamelsveld, S., & Heinemann, J. A. (2018). Agrichemicals and antibiotics in combination increase antibiotic resistance evolution. *PeerJ*, *2018*(10), 1–20. https://doi.org/10.7717/peerj.5801

Kurenbach, B., Marjoshi, D., Amábile-Cuevas, C. F., Ferguson, G. C., Godsoe, W., Paddy Gibson, A., & Heinemann, J. A. (2015). Sublethal Exposure to Commercial Formulations of the Herbicides Dicamba, 2,4-Dichlorophenoxyacetic Acid, and Glyphosate Cause Changes in Antibiotic Susceptibility in Escherichia coli and Salmonella enterica serovar Typhimurium. *MBio*, *6*(2), 1–9. https://doi.org/10.1128/mBio.00009-15

Langenheder, S., Lindström, E. S., & Tranvik, L. J. (2006). Structure and function of bacterial communities emerging from different sources under identical conditions. *Applied and Environmental Microbiology*, *72*(1), 212–220. https://doi.org/10.1128/AEM.72.1.212-220.2006

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, *9*(4), 357–359. https://doi.org/10.1038/nmeth.1923

Legendre, P., & Gallagher, E. D. (2001). Ecologically meaningful transformations for ordination of species data. *Oecologia*, *129*, 271–280. https://doi.org/10.1007/s004420100716

Leino, L., Tall, T., Helander, M., Saloniemi, I., Saikkonen, K., Ruuskanen, S., & Puigbò, P. (2020). Classification of the glyphosate target enzyme (5-enolpyruvylshikimate-3-phosphate synthase) for assessing sensitivity of organisms to the herbicide. *Journal of Hazardous Materials*, *408*(August 2020), 1–8.

https://doi.org/10.1016/j.jhazmat.2020.124556

Lenski, R. E. (2017). Experimental evolution and the dynamics of adaptation and genome evolution in microbial populations, 1–14. https://doi.org/10.1038/ismej.2017.69

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

Li, D., Liu, C. M., Luo, R., Sadakane, K., & Lam, T. W. (2015). MEGAHIT: An ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics*, *31*(10), 1674–1676. https://doi.org/10.1093/bioinformatics/btv033

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., … Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, *25*(16), 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Liao, H., Li, X., Yang, Q., Bai, Y., Cui, P., Wen, C., … Zhu, Y.-G. (2021). Herbicide Selection Promotes Antibiotic Resistance in Soil Microbiomes. *Molecular Biology and Evolution*, *38*(6), 2337–2350. https://doi.org/10.1093/molbev/msab029

Liu, C. M., McLean, P. A., Sookdeo, C. C., & Cannon, F. C. (1991). Degradation of the herbicide glyphosate by members of the family Rhizobiaceae. *Applied and Environmental Microbiology*, *57*(6), 1799–1804.

Louca, S., Jacques, S. M. S., Pires, A. P. F., Leal, J. S., Srivastava, D. S., Parfrey, L. W., … Doebeli, M. (2016). High taxonomic variability despite stable functional structure across microbial communities. *Nature Publishing Group*, *1*(December), 1–12. https://doi.org/10.1038/s41559-016-0015

Louca, S., Polz, M. F., Mazel, F., Albright, M. B. N., Huber, J. A., Connor, M. I. O., … Parfrey, L. W. (2018). Function and functional redundancy in microbial systems. *Nature Ecology & Evolution*, *2*, 936–943. https://doi.org/10.1038/s41559-018-0519-1

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, *15*(550). https://doi.org/10.1186/s13059-014-0550-8

Lozano, R. B., & Pratt, J. R. (1994). Interaction of toxicants and communities: The role

of nutrients. *Environmental Toxicology and Chemistry*, *13*(3), 361–368. https://doi.org/10.1002/etc.5620130302

Lozupone, C., & Knight, R. (2005). UniFrac : a New Phylogenetic Method for Comparing Microbial Communities. *Applied and Environmental Microbiology*, *71*(12), 8228–8235. https://doi.org/10.1128/AEM.71.12.8228

Lu, T., Xu, N., Zhang, Q., Zhang, Z., Debognies, A., Zhou, Z., … Qian, H. (2020). Understanding the influence of glyphosate on the structure and function of freshwater microbial community in a microcosm. *Environmental Pollution*, *260*(114012), 1–9. https://doi.org/10.1016/j.envpol.2020.114012

M. Oziolor, E., De Schamphelaere, K., & Matson, C. W. (2016). Evolutionary toxicology: Meta-analysis of evolutionary events in response to chemical stressors. *Ecotoxicology*, *25*(10), 1858–1866. https://doi.org/10.1007/s10646-016-1735-6

Malaj, E., Von Der Ohe, P. C., Grote, M., Kühne, R., Mondy, C. P., Usseglio-Polatera, P., … Schäfer, R. B. (2014). Organic chemicals jeopardize the health of freshwater ecosystems on the continental scale. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(26), 9549–9554. https://doi.org/10.1073/pnas.1321082111

Malley, D. F., & Mills, K. H. (1992). Whole-lake experimentation as a tool to assess ecosystem health, response to stress and recovery: the Experimental Lakes Area experience. *Journal of Aquatic Ecosystem Health*, *1*(3), 159–174. https://doi.org/10.1007/BF00044713

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.Journal*, *17*, 10–12.

Martiny, J. B. H., Jones, S. E., Lennon, J. T., & Martiny, A. C. (2015). Microbiomes in light of traits: a phylogenetic perspective. *Science*, *350*(6261), aac9323. https://doi.org/10.1126/science.aac9323

Mazor, T., Doropoulos, C., Schwarzmueller, F., Gladish, D. W., Kumaran, N., Merkel, K., … Gagic, V. (2018). Global mismatch of policy and research on drivers of biodiversity loss. *Nature Ecology and Evolution*, *2*(7), 1071–1074. https://doi.org/10.1038/s41559-018-0563-x

McComb, A. J. (2002). *Limnological Analyses*. *Lakes and Reservoirs: Research and*

Management (Vol. 7). https://doi.org/10.1046/j.1440-1770.2002.01722.x

McKnight, D. T., Huerlimann, R., Bower, D. S., Schwarzkopf, L., Alford, R. A., & Zenger, K. R. (2019). Methods for normalizing microbiome data: An ecological perspective. *Methods in Ecology and Evolution*, *10*(3), 389–400. https://doi.org/10.1111/2041-210X.13115

McMurdie, P. J., & Holmes, S. (2013). phyloseq : An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLOS One*, *8*(4), e61217. https://doi.org/10.1371/journal.pone.0061217

McMurdie, P. J., & Holmes, S. (2014). Waste not, want not : why rarefying Microbiome data is inadmissible. *PLOS Computational Biology*, *10*(4), e1003531. https://doi.org/10.1371/journal.pcbi.1003531

Mesnage, R., & Antoniou, M. N. (2018). Ignoring Adjuvant Toxicity Falsifies the Safety Profile of Commercial Pesticides. *Frontiers in Public Health*, *5*(January), 1–8. https://doi.org/10.3389/fpubh.2017.00361

Mondav, R., Bertilsson, S., Buck, M., Langenheder, S., Lindström, E. S., & Garcia, S. L. (2020). Streamlined and Abundant Bacterioplankton Thrive in Functional Cohorts. *MSystems*, *5*(5), 1–21. https://doi.org/10.1128/mSystems.00316-20

Morrissey, C. A., Mineau, P., Devries, J. H., Sanchez-Bayo, F., Liess, M., Cavallaro, M. C., & Liber, K. (2015). Neonicotinoid contamination of global surface waters and associated risk to aquatic invertebrates: A review. *Environment International*, *74*, 291–303. https://doi.org/10.1016/j.envint.2014.10.024

Mortensen, D. A., Egan, J. F., Maxwell, B. D., Ryan, M. R., & Smith, R. G. (2012). Navigating a critical juncture for sustainable weed management. *BioScience*, *62*(1), 75–84. https://doi.org/10.1525/bio.2012.62.1.12

Moss, B. (2008). Water pollution by agriculture. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *363*(1491), 659–666. https://doi.org/10.1098/rstb.2007.2176

Motta, E. V. S., Raymann, K., & Moran, N. A. (2018). Glyphosate perturbs the gut microbiota of honey bees. *Proceedings of the National Academy of Sciences*, *2018*, 201803880. https://doi.org/10.1073/pnas.1803880115

Murray, A. K., Zhang, L., Snape, J., & Gaze, W. H. (2019). Comparing the selective and

co-selective effects of different antimicrobials in bacterial communities. *International Journal of Antimicrobial Agents*, *53*(6), 767–773. https://doi.org/10.1016/j.ijantimicag.2019.03.001

Muturi, E. J., Donthu, R. K., Fields, C. J., Moise, I. K., & Kim, C. (2017). Effect of pesticides on microbial communities in container aquatic habitats. *Scientific Reports*, (November 2016), 1–10. https://doi.org/10.1038/srep44565

Nei, M., Suzuki, Y., & Nozawa, M. (2010). The Neutral Theory of Molecular Evolution in the Genomic Era. *Annu. Rev. Genomics Hum. Genet.* https://doi.org/10.1146/annurev-genom-082908-150129

Newman, M. M., Hoilett, N., Lorenz, N., Dick, R. P., Liles, M. R., Ramsier, C., & Kloepper, J. W. (2016). Glyphosate effects on soil rhizosphere-associated bacterial communities. *Science of the Total Environment*, *543*, 155–160. https://doi.org/10.1016/j.scitotenv.2015.11.008

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

Oksanen, A. J., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., Hara, R. B. O., … Wagner, H. (2018). Vegan: community ecology package. R package version 2.4-6. https://doi.org/ISBN 0-387-95457-0

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., Mcglinn, D., … Maintainer, H. W. (2019). vegan: Community Ecology Package. R package version 2.5-5. https://CRAN.R-project.org/package=vegan. *Community Ecology Package*, *2*(9), 1–297. Retrieved from https://cran.r-project.org/web/packages/vegan/vegan.pdf

Olm, M. R., Crits-Christoph, A., Bouma-Gregson, K., Firek, B. A., Morowitz, M. J., & Banfield, J. F. (2021). inStrain profiles population microdiversity from metagenomic data and sensitively detects shared microbial strains. *Nature Biotechnology*, *39*(6), 727–736. https://doi.org/10.1038/s41587-020-00797-0

Overbeek, R., Olson, R., Pusch, G. D., Olsen, G. J., Davis, J. J., Disz, T., … Stevens, R. (2014). The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). *Nucleic Acids Research*, *42*(D1), 206–214.

https://doi.org/10.1093/nar/gkt1226

Pacheco, F. S., Roland, F., & Downing, J. A. (2014). Eutrophication reverses whole-lake carbon budgets. *Inland Waters*, *4*(1), 41–48. https://doi.org/10.5268/IW-4.1.614

Paerl, H. W., Otten, T. G., & Kudela, R. (2018). Mitigating the Expansion of Harmful Algal Blooms Across the Freshwater-to-Marine Continuum. *Environmental Science and Technology*, *52*(10), 5519–5529. https://doi.org/10.1021/acs.est.7b05950

Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P., & Tyson, G. W. (2015). CheckM: Assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Research*, *25*(7), 1043–1055. https://doi.org/10.1101/gr.186072.114

Parmesan, C., & Yohe, G. (2003). A globally coherent fingerprint of climate change impacts across natural systems. *Nature*, *421*(6918), 37–42. https://doi.org/10.1038/nature01286

Patton, C. J., & Kryskalla, J. R. (2003). *Methods of Analysis by the U.S. Geological Survey National Water Quality Laboratory—Evaluation of Alkaline Persulfate Digestion as an Alternative to Kjeldahl Digestion for Determination of Total and Dissolved Nitrogen and Phosphorus in Water*. *USGS*.

Pellow, D., Mizrahi, I., & Shamir, R. (2020). PlasClass improves plasmid sequence classification. *PLoS Computational Biology*, *16*(4), 1–9. https://doi.org/10.1371/journal.pcbi.1007781

Peter, H., Beier, S., Bertilsson, S., Lindström, E. S., Langenheder, S., & Tranvik, L. J. (2011). Function-specific response to depletion of microbial diversity. *The ISME Journal*, *5*(2), 351–361. https://doi.org/10.1038/ismej.2010.119

Piggott, J. J., Niyogi, D. K., Townsend, C. R., & Matthaei, C. D. (2015). Multiple stressors and stream ecosystem functioning: Climate warming and agricultural stressors interact to affect processing of organic matter. *Journal of Applied Ecology*, *52*(5), 1126–1134. https://doi.org/10.1111/1365-2664.12480

Pimentel, D., & Pimentel, M. (1990). Comment : Adverse Environmental Consequences of the Green Revolution. *Future Options*, *16*, 329–332.

Pline-Srnic, W. (2006). Physiological Mechanisms of Glyphosate Resistance. *Weed Technology*, *20*(2), 290–300. https://doi.org/10.1614/WT-04-131R.1

Pollegioni, L., Schonbrunn, E., & Siehl, D. (2011). Molecular basis of glyphosate resistance: Different approaches through protein engineering. *The FEBS Journal*, *278*(16), 2753–2766. https://doi.org/10.1111/j.1742-4658.2011.08214.x

Preheim, S. P., Perrotta, A. R., Martin-platero, A. M., Gupta, A., & Alm, E. J. (2013). Distribution-Based Clustering : Using Ecology To Refine the Operational Taxonomic Unit. *Applied and Environmental Microbiology*, *79*(21), 6593–6603. https://doi.org/10.1128/AEM.00342-13

Quince, C., Walker, A. W., Simpson, J. T., Loman, N. J., & Segata, N. (2017). Shotgun metagenomics, from sampling to analysis. *Nature Biotechnology*, *35*(9), 833–844. https://doi.org/10.1038/nbt.3935

R Core Team. (2020). R: A Language and Environment for Statistical Computing. Vienna, Austria. Retrieved from https://www.r-project.org/

Rainio, M. J., Ruuskanen, S., Helander, M., Saikkonen, K., Saloniemi, I., & Puigbò, P. (2021). Adaptation of bacteria to glyphosate: a microevolutionary perspective of the enzyme 5-enolpyruvylshikimate-3-phosphate synthase. *Environmental Microbiology Reports*, *13*(3), 309–316. https://doi.org/10.1111/1758-2229.12931

Ramakrishnan, B., Venkateswarlu, K., Sethunathan, N., & Megharaj, M. (2019). Local applications but global implications: Can pesticides drive microorganisms to develop antimicrobial resistance? *Science of the Total Environment*, *654*, 177–189. https://doi.org/10.1016/j.scitotenv.2018.11.041

Reis, P. C. J., Thottathil, S. D., & Prairie, Y. T. (2022). The role of methanotrophy in the microbial carbon metabolism of temperate lakes. *Nature Communications*, *13*(1), 1–9. https://doi.org/10.1038/s41467-021-27718-2

Relyea, R. A. (2009). A cocktail of contaminants: How mixtures of pesticides at low concentrations affect aquatic communities. *Oecologia*, *159*(2), 363–376. https://doi.org/10.1007/s00442-008-1213-9

Rho, M., Tang, H., & Ye, Y. (2010). FragGeneScan: Predicting genes in short and error-prone reads. *Nucleic Acids Research*, *38*(20), 1–12. https://doi.org/10.1093/nar/gkq747

Roberts, T., & Hutson, D. (1999). *Metabolic Pathways of Agrochemicals. Part 2: Insecticides and Fungicides*. (T. R. Roberts, D. H. Hutson, P. W. Lee, P. H.

Nicholls, & J. R. Plimmer, Eds.), *The Royal Society of Chemistry*. Cambridge: The Royal Society of Chemistry. https://doi.org/10.1017/CBO9781107415324.004

Rocha, E. P. C., Smith, J. M., Hurst, L. D., Holden, M. T. G., Cooper, J. E., Smith, N. H., & Feil, E. J. (2006). Comparisons of dN/dS are time dependent for closely related bacterial genomes. *Journal of Theoretical Biology*, *239*(2), 226–235. https://doi.org/10.1016/j.jtbi.2005.08.037

Rodriguez-Valera, F., Martin-Cuadrado, A.-B., Rodriguez-Brito, B., Pasić, L., Thingstad, T. F., Rohwer, F., & Mira, A. (2009). Explaining microbial population genomics through phage predation. *Nature Reviews Microbiology*, *7*(11), 828–836. https://doi.org/10.1038/nrmicro2235

Rohwer, R. R., Hamilton, J. J., Newton, R. J., & McMahon, K. D. (2018). TaxAss: Leveraging a Custom Freshwater Database Achieves Fine-Scale Taxonomic Resolution. *MSphere*, *3*(5), 1–14. https://doi.org/10.1128/msphere.00327-18

Romero, F., Acuña, V., & Sabater, S. (2020). Multiple stressors determine community structure and estimated function of river biofilm bacteria. *Applied and Environmental Microbiology*, *86*(12), 1–13. https://doi.org/10.1128/AEM.00291-20

Ruiz-González, C., Archambault, E., Laforest-Lapointe, I., Del Giorgio, P. A., Kembel, S. W., Messier, C., … Beisner, B. E. (2018). Soils associated to different tree communities do not elicit predictable responses in lake bacterial community structure and function. *FEMS Microbiology Ecology*, *94*(8), 1–15. https://doi.org/10.1093/femsec/fiy115

Ruiz-González, C., Niño-García, J. P., Lapierre, J.-F., & del Giorgio, P. A. (2015). The quality of organic matter shapes the functional biogeography of bacterioplankton across boreal freshwater ecosystems. *Global Ecology and Biogeography*, *24*, 1487–1498. https://doi.org/10.1111/geb.12356

Scavia, D., David Allan, J., Arend, K. K., Bartell, S., Beletsky, D., Bosch, N. S., … Zhou, Y. (2014). Assessing and addressing the re-eutrophication of Lake Erie: Central basin hypoxia. *Journal of Great Lakes Research*, *40*(2), 226–246. https://doi.org/10.1016/j.jglr.2014.02.004

Schindler, D. W. (1998). Replication versus realism: The need for ecosystem-scale experiments. *Ecosystems*, *1*(4), 323–334. https://doi.org/10.1007/s100219900026

Schloss, P. D. (2021). Amplicon Sequence Variants Artificially Split Bacterial Genomes into Separate Clusters. *MSphere*, *6*(4), 2–7. https://doi.org/10.1128/msphere.00191-21

Shade, A., Peter, H., Allison, S. D., Baho, D. L., Berga, M., Bürgmann, H., … Handelsman, J. (2012). Fundamentals of microbial community resistance and resilience. *Frontiers in Microbiology*, *3*(DEC), 1–19. https://doi.org/10.3389/fmicb.2012.00417

Shapiro, B. J. (2016). How clonal are bacteria over time? *Current Opinion in Microbiology*, *31*, 116–123. https://doi.org/10.1016/j.mib.2016.03.013

Shapiro, B. J., Leducq, J.-B., & Mallet, J. (2016). What Is Speciation? *PLOS Genetics*, *12*(3), e1005860. https://doi.org/10.1371/journal.pgen.1005860

Shapiro, B. J., & Polz, M. F. (2014). Ordering microbial diversity into ecologically and genetically cohesive units. *Trends Microbiol. 2014*, *22*(5), 235–247. https://doi.org/10.1016/j.pestbp.2011.02.012.Investigations

Shendure, J., & Ji, H. (2008). Next-generation DNA sequencing. *Nature Biotechnology*, *26*(10), 1135–1145. https://doi.org/10.1038/nbt1486

Simon-Delso, N., Amaral-Rogers, V., Belzunces, L. P., Bonmatin, J. M., Chagnon, M., Downs, C., … Wiemers, M. (2015). Systemic insecticides (Neonicotinoids and fipronil): Trends, uses, mode of action and metabolites. *Environmental Science and Pollution Research*, *22*(1), 5–34. https://doi.org/10.1007/s11356-014-3470-y

Simpson, M. G. L. (2019). permute: Functions for Generating Restricted Permutations of Data. R package version 0.9-5. Retrieved from https://cran.r-project.org/package=permute

Singh, R. K., & Prasad, M. (2016). Advances in Agrobacterium tumefaciens-mediated genetic transformation of graminaceous crops. *Protoplasma*, *253*(3), 691–707. https://doi.org/10.1007/s00709-015-0905-3

Smedbol, É., Lucotte, M., Labrecque, M., Lepage, L., & Juneau, P. (2017). Phytoplankton growth and PSII efficiency sensitivity to a glyphosate-based herbicide (Factor 540®). *Aquatic Toxicology*, *192*(September), 265–273. https://doi.org/10.1016/j.aquatox.2017.09.021

Smith, D. P., & Peay, K. G. (2014). Sequence depth, not PCR replication, improves

ecological inference from next generation DNA sequencing. *PLoS ONE*, *9*(2). https://doi.org/10.1371/journal.pone.0090234

Smith, V. H., Joye, S. B., & Howarth, R. W. (2006). Eutrophication of freshwater and marine ecosystems. *Limnology and Oceanography*, *51*(1), 351–355.

Smith, V. H., & Schindler, D. W. (2009). Eutrophication science: where do we go from here? *Trends in Ecology and Evolution*, *24*(4), 201–207. https://doi.org/10.1016/j.tree.2008.11.009

Song, X. P., Hansen, M. C., Stehman, S. V., Potapov, P. V., Tyukavina, A., Vermote, E. F., & Townshend, J. R. (2018). Global land change from 1982 to 2016. *Nature*, *560*(7720), 639–643. https://doi.org/10.1038/s41586-018-0411-9

Sorek, R., Kunin, V., & Hugenholtz, P. (2008). CRISPR--a widespread system that provides acquired resistance against phages in bacteria and archaea. *Nature Reviews. Microbiology*, *6*(3), 181–186. https://doi.org/10.1038/nrmicro1793

Springmann, M., Clark, M., Mason-D'Croz, D., Wiebe, K., Bodirsky, B. L., Lassaletta, L., … Willett, W. (2018). Options for keeping the food system within environmental limits. *Nature*, *562*(7728), 519–525. https://doi.org/10.1038/s41586-018-0594-0

Stachowski-Haberkorn, S., Becker, B., Marie, D., Haberkorn, H., Coroller, L., & de la Broise, D. (2008). Impact of Roundup on the marine microbial community, as shown by an in situ microcosm experiment. *Aquatic Toxicology*, *89*(4), 232–241. https://doi.org/10.1016/j.aquatox.2008.07.004

Starr, A. V., Bargu, S., Maiti, K., & DeLaune, R. D. (2017). The Effect of Atrazine on Louisiana Gulf Coast Estuarine Phytoplankton. *Archives of Environmental Contamination and Toxicology*, *72*(2), 178–188. https://doi.org/10.1007/s00244-016-0335-z

Staub, J. M., Brand, L., Tran, M., Kong, Y., & Rogers, S. G. (2012). Bacterial glyphosate resistance conferred by overexpression of an E. coli membrane eZux transporter. *Journal of Industrial Microbiology and Biotechnology*, *39*(4), 641–647. https://doi.org/10.1007/s10295-011-1057-x

Steffen, W., Crutzen, P. J., & McNeill, J. R. (2007). The anthropocene: Are humans now overwhelming the great forces of nature? *Ambio*, *36*(8), 614–621. https://doi.org/10.1579/0044-7447(2007)36[614:TAAHNO]2.0.CO;2

Stehle, S., & Schulz, R. (2015). Agricultural insecticides threaten surface waters at the global scale. *Proceedings of the National Academy of Sciences of the USA*, *112*(18), 5750–5755. https://doi.org/10.1073/pnas.1500232112

Tenaillon, O., Barrick, J. E., Ribeck, N., Deatherage, D. E., Blanchard, J. L., Dasgupta, A., … Lenski, R. E. (2016). Tempo and mode of genome evolution in a 50,000-generation experiment. *Nature*, *536*(7615), 165–170. https://doi.org/10.1038/nature18959

Thibodeau, G., Walsh, D. A., & Beisner, B. E. (2015). Rapid eco-evolutionary responses in perturbed phytoplankton communities. *Proceedings of the Royal Society B*, *282*, 20151215. https://doi.org/10.1098/rspb.2015.1215

Thompson, M. S. A., Bankier, C., Bell, T., Dumbrell, A. J., Gray, C., Ledger, M. E., … Woodward, G. (2016). Gene-to-ecosystem impacts of a catastrophic pesticide spill: testing a multilevel bioassessment approach in a river ecosystem. *Freshwater Biology*, *61*(12), 2037–2050. https://doi.org/10.1111/fwb.12676

Tilman, D., Fargione, J., Wolff, B., D'Antonio, C., Dobson, A., Howarth, R., … Swackhamer, D. (2001). Forecasting agriculturally driven global environmental change. *Science*, *292*(5515), 281–284. https://doi.org/10.1126/science.1057544

Tollefson, J. (2019). Brazil debates loosening environmental protections. *Nature*, *539*(4), 147–148.

Tranvik, L. J., Downing, J. A., Cotner, J. B., Loiselle, S. A., Striegl, R. G., Ballatore, T. J., … Weyhenmeyer, G. A. (2009). Lakes and reservoirs as regulators of carbon cycling and climate, *54*(1), 2298–2314.

Tromas, N., Fortin, N., Bedrani, L., Terrat, Y., Cardoso, P., Bird, D., … Shapiro, B. J. (2017). *Characterising and predicting cyanobacterial blooms in an 8-year amplicon sequencing time course*. *The ISME Journal* (Vol. 11). https://doi.org/10.1038/ismej.2017.58

Tsui, M. T. K., & Chu, L. M. (2003). Aquatic toxicity of glyphosate-based formulations: Comparison between different organisms and the effects of environmental factors. *Chemosphere*, *52*(7), 1189–1197. https://doi.org/10.1016/S0045-6535(03)00306-0

van Bruggen, A. H. C., He, M. M., Shin, K., Mai, V., Jeong, K. C., Finchk, M. R., & Morris, J. J. (2018). Environmental and health effects of the herbicide glyphosate.

Science of The Total Environment, *617*, 255–268. https://doi.org/10.1016/j.scitotenv.2017.10.309

Van den Brink, P. J., den Besten, P. J., bij de Vaate, A., & ter Braak, C. J. F. (2009). Principal response curves technique for the analysis of multivariate biomonitoring time series. *Environmental Monitoring and Assessment*, *152*, 271–281. https://doi.org/10.1007/s10661-008-0314-6

Van den Brink, P. J., & Ter Braak, C. J. F. (1999). Principal response curves: Analysis of time-dependent multivariate responses of biological community to stress. *Environmental Toxicology and Chemistry*, *18*(2), 138–148. https://doi.org/10.1002/etc.5620180207

Van Rij, J., Wieling, M., Baayen, H. R., & van Rijn, H. (2020). Interpreting Time Series and Autocorrelated Data Using GAMMs Author.

Vörösmarty, C. J., McIntyre, P. B., Gessner, M. O., Dudgeon, D., Prusevich, A., Green, P., … Davies, P. M. (2010). Global threats to human water security and river biodiversity. *Nature*, *467*(7315), 555–561. https://doi.org/10.1038/nature09440

Wang, C., Lin, X., Li, L., Lin, L., & Lin, S. (2017). Glyphosate Shapes a Dinoflagellate-Associated Bacterial Community While Supporting Algal Growth as Sole Phosphorus Source. *Frontiers in Microbiology*, *8*(2530), 1–13. https://doi.org/10.3389/fmicb.2017.02530

Wang, Y., Lu, J., Engelstädter, J., Zhang, S., Ding, P., Mao, L., … Guo, J. (2020). Non-antibiotic pharmaceuticals enhance the transmission of exogenous antibiotic resistance genes through bacterial transformation. *ISME Journal*, 2179–2196. https://doi.org/10.1038/s41396-020-0679-2

Weiss, S., Xu, Z. Z., Peddada, S., Amir, A., Bittinger, K., Gonzalez, A., … Knight, R. (2017). Normalization and microbial differential abundance strategies depend upon data characteristics. *Microbiome*, *5*(1), 27. https://doi.org/10.1186/s40168-017-0237-y

Wetzel, R. G., & Likens, G. E. (2000). *Limnological Analyses*. *Springer Science & Business Media* (3rd ed., Vol. 7). New York, New York, USA. https://doi.org/10.1046/j.1440-1770.2002.01722.x

Wickham, H. (2009). ggplot2: Elegant Graphics for Data Analysis. *Media*, *35*(July), 211.

https://doi.org/10.1007/978-0-387-98141-3

Wiedenbeck, J., & Cohan, F. M. (2011). Origins of bacterial diversity through horizontal genetic transfer and adaptation to new ecological niches. *FEMS Microbiology Reviews*, *35*(5), 957–976. https://doi.org/10.1111/j.1574-6976.2011.00292.x

Williamson, C. E., Dodds, W., Kratz, T. K., & Palmer, M. A. (2008). Lakes and streams as sentinels of environmental change in terrestrial and atmospheric processes. *Frontiers in Ecology and the Environment*, *6*(5), 247–254. https://doi.org/10.1890/070140

Williamson, C. E., Saros, J. E., Vincent, W. F., & Smol, J. P. (2009). Lakes and reservoirs as sentinels, integrators, and regulators of climate change. *Limnology and Oceanography*, *54*(6 PART 2), 2273–2282. https://doi.org/10.4319/lo.2009.54.6_part_2.2273

Wittmer, I. K., Bader, H. P., Scheidegger, R., Singer, H., Lück, A., Hanke, I., … Stamm, C. (2010). Significance of urban and agricultural land use for biocide and pesticide dynamics in surface waters. *Water Research*, *44*(9), 2850–2862. https://doi.org/10.1016/j.watres.2010.01.030

Wood, S. (2017). *Generalized Additive Models* (2nd ed.). New York: Chapman and Hall/CRC. Retrieved from https://doi.org/10.1201/9781315370279

Wright, E. S. (2016). Using DECIPHER v2.0 to analyze big biological sequence data in R. *The R Journal*, *8*(1), 352–359. https://doi.org/10.32614/rj-2016-025

Xing, Y., Wu, S., & Men, Y. (2020). Exposure to Environmental Levels of Pesticides Stimulates and Diversifies Evolution in Escherichia coli toward Higher Antibiotic Resistance. *Environmental Science & Technology*. https://doi.org/10.1021/acs.est.0c01155

Yamamuro, M., Komuro, T., Kamiya, H., Kato, T., Hasegawa, H., & Kameda, Y. (2019). Neonicotinoids disrupt aquatic food webs and decrease fishery yields. *Science*, *366*(6465), 620–623. https://doi.org/10.1126/science.aax3442

Zalasiewicz, J., Williams, M., Haywood, A., & Ellis, M. (2011). The anthropocene: A new epoch of geological time? *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *369*(1938), 835–841. https://doi.org/10.1098/rsta.2010.0339

Zhang, H., Liu, J., Wang, L., & Zhai, Z. (2021). Glyphosate escalates horizontal transfer of conjugative plasmid harboring antibiotic resistance genes. *Bioengineered*, *12*(1), 63–69. https://doi.org/10.1080/21655979.2020.1862995

Zhang, Z., Qu, Y., Li, S., Feng, K., Wang, S., Cai, W., … Deng, Y. (2017). Soil bacterial quantification approaches coupling with relative abundances reflecting the changes of taxa. *Scientific Reports*, *7*(1), 1–11. https://doi.org/10.1038/s41598-017-05260-w

Zhu, Y. G., & Penuelas, J. (2020). Changes in the environmental microbiome in the Anthropocene. *Global Change Biology*, *26*(6), 3175–3177. https://doi.org/10.1111/gcb.15086