TURUN
YLIOPISTO
UNIVERSITY
OF TURKU

Investigating microbial communities

Current Approaches and Future Directions

Niina Smolander

TURUN YLIOPISTON JULKAISUJA – ANNALES UNIVERSITATIS TURKUENSIS
SARJA – SER. AII OSA – TOM. 425 | BIOLOGICA – GEOGRAPHICA – GEOLOGICA | TURKU 2026

TURUN
YLIOPISTO
UNIVERSITY
OF TURKU

INVESTIGATING MICROBIAL COMMUNITIES

Current Approaches and Future Directions

Niina Smolander

TURUN YLIOPISTON JULKAISUJA – ANNALES UNIVERSITATIS TURKUENSIS
SARJA – SER. AII OSA – TOM. 425 | BIOLOGICA – GEOGRAPHICA – GEOLOGICA | TURKU 2026

University of Turku

Faculty of Science
Department of Biology
Biology
Doctoral programme in Biology, Geography and Geology (BGG)

Supervised by

Docent, Manu Tamminen
Department of Biology
University of Turku
Turku, Finland

Professor, Teppo Hiltunen
Department of Biology
University of Turku
Turku, Finland

Reviewed by

Professor, Ellen Decaestecker
Department of Biology
KU Leuven
Leuven, Belgium

Docent, Derek Lundberg
Department of Plant Biology
Swedish University of Agricultural Sciences
Uppsala, Sweden

Opponent

Associate Professor, Lars Behrendt
Department of Organismal Biology
Uppsala University
Uppsala, Sweden

ISBN 978-952-02-0501-0 (PRINT)
ISBN 978-952-02-0502-7 (PDF)
ISSN 0082-6979 (Print)
ISSN 2343-3183 (Online)
Painosalama, Turku, Finland 2026

UNIVERSITY OF TURKU
Faculty of Science
Department of Biology
Biology
NIINA SMOLANDER: Investigating microbial communities – Current Approaches and Future Directions
Doctoral Dissertation, 123 pp.
Doctoral Programme in Biology, Geography and Geology (BGG)
January 2026

ABSTRACT

Bacteria form heterogeneous communities with specific spatial structures. The heterogeneity of bacteria can be genetic or phenotypic and even within one strain of bacteria there can be differences in the metabolism and behavior between the cells. The most common bulk method, the amplification and sequencing of 1–3 hypervariable regions of the 16S rRNA gene, reveals the assemblage of the bacterial community, from which the functional potential of the community can be inferred. In this thesis, I used 16S rRNA gene sequencing to examine how two commonly used agrochemicals, glyphosate-based herbicide (GBH) and phosphate, affect the endophytic bacterial communities of various potato, faba bean and oat tissues in the early vegetative and the late flowering growth stages (**Chapter I**). Changes in the community diversity and/or composition were seen in nearly all the plant tissues, in at least one of the growth stages, more so in response to phosphate than GBH. I also used the 16S rRNA gene sequencing approach to study how the gut microbiota of two odonate species, *Lestes sponsa* and *Sympetrum vulgatum,* varied in response to environmental factors, species, sex, sampling site and season, and diet (**Chapter II**). The diversity of the gut microbiota was higher in *S. vulgatum* than in *L. sponsa*, potentially in response to a more diverse diet. The diet affected the microbiota assemblages more than the environmental factors, but the effect was not strong, potentially due to large variation between individuals. As the 16S rRNA gene sequencing approach and other bulk methods cannot uncover details about the heterogeneity and spatial structuring of the microbiome, novel single-cell methods are required. **Chapter III** presents Prider, a computational tool we developed for high-throughput primer and probe designing. **Chapter IV** describes a single-cell sequencing based method we developed that combines porous polyacrylamide beads and split-pool barcoding. The method does not require specialised equipment, such as microfluidic systems, and once fully optimised, can be used for the analysis of both the microbial genetic heterogeneity as well as microbial interactions in communities, expanding the kind of information that can be gained from microbiome research, including studies such as those described in **Chapter I** and **Chapter II**.

KEYWORDS: agrochemicals, barcoding, bioinformatics, microbiomes, oligonucleotides, single-cell sequencing

TURUN YLIOPISTO
Matemaattis-luonnontieteellinen tiedekunta
Biologian laitos
Biologia
NIINA SMOLANDER: Mikrobiyhteisöjen tutkimus – nykyiset lähestymistavat ja tulevaisuuden suuntaviivat
Väitöskirja, 123 s.
Biologian, maantieteen ja geologian tohtoriohjelma (BGG)
Tammikuu 2026

TIIVISTELMÄ

Bakteerit muodostavat heterogeenisiä ja rakenteellisesti järjestäytyneitä yhteisöjä. Heterogeenisyys voi olla geneettistä tai fenotyyppistä, mikä näkyy eroina mm. solujen metaboliassa ja käytöksessä. Bakteeriyhteisöjä tutkitaan yleisimmin menetelmillä, jotka perustuvat 1–3 16S rRNA -geenin hypervariaabelisen alueen monistukseen ja sekvensointiin. Näillä menetelmillä saadaan selville yhteisön koostumus, josta voidaan päätellä yhteisön ominaisuuksia. Väitöskirjassani käytin 16S rRNA -geenisekvensointia sen tutkimiseen, kuinka kaksi yleisesti maataloudessa käytettyä kemikaalia, rikkakasvien torjunta-aine glyfosaatti ja lannoite fosfaatti, vaikuttavat endofyyttisiin bakteeriyhteisöihin perunan, kauran ja härkäpavun eri osissa varhaisessa kasvuvaiheessa ja kukkimisvaiheessa (**Luku I**). Muutoksia yhteisöjen diversiteetissä sekä koostumuksessa havaittiin lähes kaikissa tutkituissa kasvin osissa, ainakin yhdessä vaiheista, useammin fosfaatin kuin glyfosaatin aiheuttamana. Käytin myös 16S rRNA -geenisekvensointia sirokeijukorennon (*Lestes sponsa*) ja punasyyskorennon (*Sympetrum vulgatum*) suolistomikrobistojen vertailuun näiden lajien, eri sukupuolten, eri asuinpaikkojen, eri ajankohtien ja ruokavalioiden välillä (**Luku II**). Punasyyskorennon mikrobiston diversiteetti oli korkeampi, mahdollisesti johtuen lajin monipuolisemmasta ravinnosta. Ravinnolla oli suurempi vaikutus suolistomikrobiston koostumukseen kuin muilla tutkituilla tekijöillä, vaikkakin vaikutus oli kaiken kaikkiaan pieni, mahdollisesti johtuen suuresta vaihtelusta yksilöiden välillä. Koska erilaisilla koko yhteisön sekvensointiin perustuvilla ”bulk”-menetelmillä, kuten 16S rRNA -geenin sekvensoinnilla, ei pystytä tarkastelemaan mikrobiomin heterogeenisyyttä tai rakennetta, tarvitaan yksisolumenetelmiä. **Luvussa III** esitellään R-paketti Prider, joka mahdollistaa suurikapasiteettisen aluke- ja koetinsuunnittelun. **Luvussa IV** kuvaillaan uusi polyakryyliamidihelmet ja split-pool viivakoodauksen yhdistävä yksisolusekvensointimenetelmä, joka ei vaadi normaalista poikkeavia laboratoriolaitteita, kuten mikrofluidistiikkaan perustuvat menetelmät. Kun menetelmä on täysin optimoitu, sitä voidaan käyttää mikrobiyhteisöjen sisäisen geneettisen heterogeenisyyden sekä vuorovaikutusten tutkimiseen, laajentaen sitä millaista tietoa **Luvuissa I** ja **II** esitetyistä yhteisöistä voidaan tulevaisuudessa saada.

ASIASANAT: maatalouskemikaalit, viivakoodaus, bioinformatiikka, mikrobiomit, oligonukleotidit, yksisolusekvensointi

Table of Contents

Abbreviations

ASV	Amplicon sequence variant
CAP	Constrained analysis of principal coordinates
EPSP	Enzyme 5-enolpyruvylshikimate-3-phosphate
FACS	Fluorescence-activated cell sorting
GBH	Glyphosate-based herbicide
HGT	Horizontal gene transfer
NGS	Next-generation sequencing
UMI	Unique molecular identifier
WGA	Whole genome amplification
ZOTU	Zero-radius operational taxonomic unit

List of Original Publications

This dissertation is based on the following original publications, which are referred to in the text by their Roman numerals:

I Smolander, N., Fuchs, B., Helander, M., Puigbò, P., Tamminen, M., Saikkonen, K. and Mathew, S. A. Glyphosate and phosphate treatments in soil differentially affect crop microbiomes depending on species, tissue and growth stage. *Scientific Reports*, 2025; 15: 25502. https://doi.org/10.1038/s41598-025-11430-y

II Kaunisto, K. M., Smolander, N., Tamminen, M., Stoks, R., Kylänpää, S., Sääksjärvi, I. E. and Vesterinen, E. J. The Invisible Web: Linking Diet and Gut Microbiota Across Time and Space in Odonates. Unpublished manuscript.

III Smolander, N., Julian, T. R. and Tamminen, M. Prider: multiplexed primer design using linearly scaling approximation of set coverage. *BMC Bioinformatics*, 2022; 23: 174. https://doi.org/10.1186/s12859-022-04710-1

IV Smolander, N., Talvitie, J. and Tamminen, M. Polyacrylamide bead split-pool method for microbial community analysis. *bioRxiv*, 2025. https://doi.org/10.1101/2025.06.26.661670

1 Introduction

1.1 Microbial communities

Microorganisms — bacteria, archaea, fungi, small protists and algae — are almost everywhere (Berg et al., 2020; Cockell, 2021). Microorganisms form multidomain microbial communities that consist of different microbes living together in a particular environment. These communities are also known as microbiota or microbiome. While the definition of microbiota only covers the assemblage of the community, the definition of microbiome also includes the "theatre of activity" of the microbes and hence can extend to non-living carriers of genetic material, such as viruses and plasmids. Microbiomes are dynamic; they vary in time and space and even within one organ, such as intestines, various sub-microbiomes may form. (Berg et al., 2020). Multicellular organisms are to some extent dependent on microbes, but the level of dependency of the microbiome varies between hosts. While some hosts, such as humans, cows and plants, are dependent on their microbiome, the others, such as ants and caterpillars, seemingly lack a stable microbiome (Hammer et al., 2019; Dastogeer et al., 2020). The microbial dependency of animals and plants is reflected in how the disruption of the microbiome affects the health of the host. For example, in humans various gastrointestinal, metabolic and neurological diseases and disorders have been associated with the dysbiosis of gut microbiota (Shen et al., 2025). In plants, dysbiosis is associated with a lack of nutrient intake, reduced stress tolerance or various diseases (Arnault et al., 2023).

The concept of microbiome goes beyond co-existing microbial cells: microbes both produce substances as well as sense their environment, which results in interactions between cells in microbiomes (Berg et al., 2020). Microbes interact with each other in a variety of ways: from mutually beneficial relationships to those that are harmful to all cells in the community, such as in the case of competition. Interactions may also be neutral, commensalistic, amensalistic or such as in the case of predation; beneficial to just one type of cell, but harmful to others. (Faust and Raes, 2012). Bacterial cells communicate with each other by secreting metabolites and small molecules as well as by altering environmental factors, such as the pH level (Pierce and Dutton, 2022). For example, quorum sensing — a process where bacteria regulate their gene expression based on the population density — is based

on small molecules called autoinducers (Miller and Bassler, 2001). Cooperative interactions emerge for example when some bacterial cells produce metabolites that other bacterial cells can in turn use as nutrients in a process called metabolic cross-feeding. These chain-like processes can be mediated by diffusion of the metabolites or more directly by intercellular nanotubes between cells. (Pande et al., 2015). Bacterial antagonistic interactions can be based on contact-dependent systems, such as the direct transfer of toxic effector proteins, or contact-independent systems, such as the secretion of antibiotics, and have an effect on the bacterial community diversity as well as the success of bacterial colonisation in a new environment, for example during infection (Fu et al., 2018; Wang et al., 2022).

In bacterial communities, beneficial interactions, such as metabolic cross-feeding, can result in physical proximity between cells, whereas disadvantageous interactions, such as competition, may result in avoidance between the competing strains or species (Nielsen et al., 2000). These kinds of cell-cell interactions, where some cells are more prone to be in the near vicinity of other bacterial cells while others actively avoid cells with certain traits such as antibiotic producers, are drivers of the spatial structure of microbiomes. Additionally, microbial interactions with their environment, active migration as well as passive dispersal all contribute to the physical structure of microbiomes. (Yanni et al., 2019). The cell-cell interactions also allow the formation of biofilms, in which bacteria adhere to each other and to a surface. The cells in biofilms differ phenotypically from independent cells. (Costerton et al., 1995). The spatial structure of the community also in return affects the way microbes interact with each other (Pande et al., 2016; Lobanov et al., 2023). For instance, biofilms can both accelerate the accumulation of antibiotic resistance mutations as well as aid the survival of those antibiotic-resistant bacteria whose mutations come with a fitness cost in a non-antibiotic environment (France et al., 2018).

Although in nature microbial communities are made up of organisms belonging to all three domains of life, the focus of this thesis is mainly on bacterial communities. The following two sections briefly discuss the plant microbiomes and insect gut microbiomes, which are the central topics of **Chapter I** and **Chapter II**, respectively. Section 1.1.3. focuses on microbial heterogeneity, which, alongside the spatial structuring of microbial communities, is the rationale behind the methodological approach taken in **Chapter IV**. The concept of microbial heterogeneity also contributes to the motivation behind **Chapter III**.

1.1.1 Plant microbiomes

Plants provide a variety of niches for microbes and microbial communities. Just within one single plant there can be various distinct endophytic and epiphytic

subcommunities that reside in and on (respectively) different plant tissues and organs, including the stem, leaves, flowers, fruits, seeds and roots (Compant et al., 2010). Microbiomes are beneficial for their host plant for a variety of reasons, such as by providing improved nutrient acquisition (Zhang et al., 2019; de Lima et al., 2024), tolerance against abiotic stress (Kaur and Karnwal, 2023) and protection against pathogenic bacteria and fungi (Carrión et al., 2019). Often plants also interact in a variety of ways with the bacterial communities residing in the soil surrounding the plant roots, known as the rhizosphere. The microbial species populating the rhizosphere, as well as their activity, differ from those of the bulk soil microbiome. (Ling et al., 2022).

Various factors impact the plant microbiome: host related factors such as the health, metabolites and developmental stage of the plant, interactions with other microbes and environmental factors such as various soil properties, the climate and agricultural practices (Dastogeer et al., 2020). Plants produce a variety of substances and exudates such as phytohormones, which function as chemoattractants for the recruitment of beneficial microbial symbionts, especially in the roots (Feng et al., 2019; Rico-Jiménez et al., 2022) and maintain the microbiome homeostasis in the phyllosphere (Pfeilmeier et al., 2021; Su et al., 2024). Environmental factors, such as agricultural practices, can affect the plant microbiome either directly or indirectly, via altered plant phenotype (Ruuskanen et al., 2023). Glyphosate-based herbicides (GBHs), one of the most commonly used herbicides (Myers et al., 2016), function by inhibiting the enzyme 5-enolpyruvylshikimate-3-phosphate (EPSP) synthase of the shikimate pathway required for the synthesis of essential aromatic amino acids in plants (Schönbrunn et al., 2001). However, bacteria are also known to utilise the same pathway and depending on the amino acid content in the active site of the EPSP synthase they produce, they may be sensitive or resistant to glyphosate and hence may be directly affected by GBHs (Rainio et al., 2021). Other widely used agrochemicals with potential direct adverse effects on plant microbiotas include the fungicides metiram and captan (Walter et al., 2007) as well as herbicides S-metolachlor (Xu et al., 2020) and imazethapyr (Zheng et al., 2022).

1.1.2 Insect gut microbiomes

Insects are vital for ecosystems: they pollinate, function as a food source for other animals, and take part in ecosystem processes such as decomposition and nutrient cycling (Losey and Vaughan, 2006; Yang and Gratton, 2014). Although the level of dependency on microbiomes seems to vary between different insect species (Hammer et al., 2019), insect microbiomes have been shown to be essential to their hosts in various cases: gut microbiomes can provide protection against pathogens (Wang and Rozen, 2018; Dosch et al., 2021), regulate the development (Coon et al.,

2020), modulate the behaviour (Sharon et al., 2011; Wong et al., 2017), aid stress-tolerance (Theys et al., 2023) and digestion (Tokuda and Watanabe, 2007) as well as provide nutrients (Sabree et al., 2009) to their insect hosts. Some of these interactions between the host insect and bacteria are obligatory, with millions of years of codiversification, such as the relationship between aphids and the bacteriome-associated gammaproteobacteria *Buchnera aphidicola*, which provide amino acids to their host and are required for normal host development (Moran et al., 2008). Some insect microbiomes can also be very dynamic. For example, insects that undergo metamorphosis, such as dragonflies and butterflies, can harbour different gut microbiotas as adults compared to their nymph or larvae stage, potentially in response to a shift in their habitat (Nobles and Jackson, 2020; Wang et al., 2020).

Both obligatory and facultative microbial symbionts can be acquired by vertical transmission, most often maternally (Moran et al., 2008). Horizontal transmission of symbionts may take place for example via soil, diet, faeces or other members of the insect community (Lange et al., 2023). Additionally, a variety of environmental factors, such as the geographical region (Li et al., 2022), temperature, and diet (Montagna et al., 2015) as well as pesticides (Helander et al., 2023; Ren et al., 2023) can affect the diversity and the assemblage of the microbiota. The maintenance of the microbiome is achieved through various mechanisms. Some obligate and facultative inheritable bacteria reside in specialised cells known as bacteriocytes, which often form an organ called bacteriome to which bacteria are in some cases selected via cellular mechanisms (Koga et al., 2012; Ferrarini et al., 2022). This kind of compartmentalisation of cells helps to maintain the microbiota by protecting the bacteria from the host immune system (Ferrarini et al., 2022). As only some microbiotas reside in bacteriome and as the insect gut constantly encounters new bacteria, some of which are pathogenic, distinguishing the gut microbiota from the harmful bacteria is important. Insects regulate the microbiome homeostasis by physical barriers and a variety of signalling pathways. Additionally, the microbiome itself regulates the conditions in the gut, aiding its own maintenance. (Bai et al., 2021).

1.1.3 Bacterial cellular heterogeneity

Bacterial communities are inherently heterogeneous as they contain various bacterial taxa. Cellular heterogeneity, which can be genetic or phenotypic, is however not limited to taxonomy or inter-species differences but can be observed even in initially clonal communities (Spudich and Koshland, 1976; Croucher et al., 2014). Understanding cellular heterogeneity is important due to its role in infectious diseases, antibiotic persistence of bacteria and the spread of antibiotic resistance (Balaban et al., 2019; Andam, 2019).

Various mechanisms can lead to genetic heterogeneity: *de novo* point mutations or indels, horizontal gene transfer (HGT), chromosomal duplications, changes in the number of mobile genetic elements, and unbalanced distribution of genetic material during cell division (Brehm-Stecher and Johnson, 2004). Useful mutations, such as the ones that confer resistance to antibiotics, appear even in environments where the selection pressure is low, e.g. the level of antibiotic is below minimal inhibitory concentration (Wistrand-Yuen et al., 2018) or non-existent (David, 1970), resulting in natural heterogeneity of the population. In some cases, the changes in the genome are reversible: in phase variation, mutations function as an on-off-switch for a specific phenotype related to features such as biofilm formation, synthesis of pili or immune evasion and occur more frequently than spontaneous mutations. Phase variation can take place as a result of numerous mechanisms including site-specific inversion and duplication. (Wisniewski-Dyé and Vial, 2008).

Phenotypic heterogeneity, exhibited by differences in biochemistry, physiology or behaviour between cells, is only partially explained by the genetic heterogeneity of cells (Brehm-Stecher and Johnson, 2004). Even though some proportion of the phenotypic heterogeneity is stochastic (Elowitz et al., 2002), some variation, for example in the metabolism of the cells, is driven by environmental conditions, such as the availability of nutrients (Schreiber et al., 2016). Additionally, some forms of heterogeneity, such as the heterogeneity of cell sporulation, can be driven by epigenetic inheritance (Veening et al., 2008). Phenotypic heterogeneity can benefit the bacterial community or individual cells in a variety of ways, such as by allowing the division of labour or by allowing some cells — also known as cheaters — to benefit from the surrounding community without contributing to it themselves. Furthermore, in a dynamic environment where the conditions can change drastically in a period of time too short for the bacteria to adapt their phenotype, some individual cells may express features, which result in lower fitness in the current environment, but allow them to survive should the conditions change. (Ackermann, 2015). Although the majority of the research of the cellular heterogeneity has been conducted *in vitro* and its significance in natural environments is mostly elusive, some evidence of its existence and importance to natural bacterial communities exists, such as for the filamentous bacterial species *Candidatus* Microthrix parvicella in wastewater treatment plants (Ackermann, 2015; Sheik et al., 2016).

1.2 Analysis of bacterial communities

Microbes and microbial communities play an essential role in agriculture, industry, and the health and well-being of humans and other organisms as well as the environment (Ma et al., 2023; Virgo et al., 2025). Yet, there are more than a thousand pathogenic microbes identified, which, albeit cover only a tiny fraction of all the

microbial species on Earth, cause a variety of infections in humans and are a major concern due to the global antimicrobial resistance crisis (Balloux and van Dorp, 2017; Bartlett et al., 2022). The ecology of microbial communities is closely connected to the maintenance and spread of antibiotic resistance and virulence factors, as well as the survival of opportunistic pathogens (Tang et al., 2021; Bottery et al., 2021; Zhuang et al., 2024). More detailed understanding of microbial communities, including their structure, function, interactions, heterogeneity and evolution, is essential for understanding both the beneficial and detrimental interactions between microbes and their surroundings. This requires methods that allow comprehensive analysis of microbiomes. Alongside laboratory methods, computational tools that support and aid the designing of the experiments, such as tools for primer design, as well as computational tools for the analysis of data, are essential (Kreer et al., 2020; Dakal et al., 2025). The following three sections focus on the laboratory methods for bulk, single cell, and spatial analyses of microbial communities.

1.2.1 Bulk approaches

Microbiotas are most commonly studied using 16S rRNA gene amplicon sequencing, which reveals the taxonomic composition of the community (Bose and Moore, 2023; Bartoš et al., 2024). The 16S rRNA gene is approximately 1,500 bp long and contains nine variable regions. Due to the limitations of widely used and cost-efficient next-sequencing approaches, the sequencing of the whole gene has not been a common practice, instead somewhere between one to three sub-regions of the gene are usually targeted, which is only accurate at the genus-level or higher. (Johnson et al., 2019). The development of accurate long-read sequencing methods however has facilitated the sequencing of the whole 16S rRNA gene, thus providing more precision in taxonomical classification, but these methods are more expensive than the sequencing of the sub-regions (Buetas et al., 2024). Although the 16S rRNA gene mainly provides taxonomic abundance information, by utilising databases, a certain degree of approximations of the functional potential and features of the community can also be made (Langille et al., 2013; Iwai et al., 2016; Mathew et al., 2022). However, as the functional information obtainable from 16S rRNA sequencing is limited, metagenomic sequencing — the sequencing of all the DNA of a sample — provides a more accurate functional profile of the community and also a better taxonomic resolution in well-established systems such as human microbiomes (Hillmann et al., 2018). A more accurate functional profile of the bacterial community can be obtained by using other “omics” approaches, such as transcriptomics, metabolomics, or proteomics, which target the whole of the RNA, metabolites or proteins of a sample, respectively. Metatranscriptomics utilises

similar sequencing-based methods as metagenomics, while metabolomics and proteomics are typically based on mass-spectrometry. (Galloway-Peña and Hanson, 2020).

1.2.2 Single-cell approaches

As microbial communities are heterogeneous and even bacteria belonging to the same taxa may differ in their genomic constitution or function, bulk approaches, such as metagenomics and metatranscriptomics, are not sufficient enough for capturing the heterogeneity. Hence, single-cell approaches are needed. (Brehm-Stecher and Johnson, 2004). Additionally, as microbial communities can be extremely complex (Leviatan et al., 2022; Labouyrie et al., 2023) high-throughput methods are necessary for more detailed understanding of their assemblage and function (Tedersoo et al., 2021; Stromberg et al., 2023). High-throughput analysis is achieved through next-generation sequencing (NGS) methods (Koboldt et al., 2013).

Single-cell approaches usually require some form of compartmentalisation of individual cells, which can be achieved using a variety of methods such as fluorescence-activated cell sorting (FACS) or droplet microfluidics (Lloréns-Rico et al., 2022). As the compartmentalisation in FACS is performed using multiwell plates, it is not suitable for high-throughput sequencing of highly heterogeneous samples, due to the limitations in the number of cells that can realistically be sorted and the number of subsequent reactions required for the amplification of the nucleic acid (Zilionis et al., 2016). One potential solution for making FACS more cost-effective and high-throughput is to perform the amplification in a smaller reaction volume as suggested by Sobol and Kaster (2023).

In droplet microfluidics, the compartments are nano or picolitre-sized aqueous droplets, which are formed using a microfluidic device and are separated by a continuous phase, such as oil (Lloréns-Rico et al., 2022). The aqueous phase can also be for instance solidified using agarose (Lan et al., 2017) or be turned into semi-permeable capsules (Leonaviciene et al., 2020). Various droplet microfluidics-based single-cell methods for studying both the genetic and phenotypic (transcriptional) heterogeneity have emerged during the last decade. For example, SiC-seq (single-cell genomic sequencing), a method based on the barcoding of fragmented DNA trapped in agarose beads, has been utilised for the analysis of seawater microbial communities (Lan et al., 2017). Another example is SAG-gel (single-cell amplified genomes in gel beads), a method based on agarose beads and whole genome amplification (WGA) that has been used for studying bacterial genetic heterogeneity in soil, seawater and human skin (Ide et al., 2022; Nishikawa et al., 2022). In scALA (single-cell amplified genome long-read assembly) SAG-gel was combined with long-read sequencing to improve the completeness of the coverage for the analysis

of human gut microbiomes (Kogawa et al., 2023). To analyse single-cell transcriptomes, advanced droplet microfluidics methods, such as M3-seq (massively-parallel, multiplexed, microbial sequencing) by Wang et al., (2023), and ProBac-seq (probe-based bacterial sequencing) by McNulty et al. (2023), have been developed. M3-seq has been utilised for identifying bacterial subpopulations and phage infection, while ProBac-seq has been used to study heterogeneity in transcriptional states related to toxin expression, metabolic pathways and physiology.

As droplet microfluidics often require specialised equipment, they are not accessible by all laboratories (Elvira et al., 2022; Sobol and Kaster, 2023) and hence, other approaches for separating the cells from one another have been developed. In epicPCR (emulsion, paired isolation and concatenation PCR), the compartmentalisation is achieved by emulsifying a cell-acrylamide-oil mixture using a vortex and allowing the acrylamide droplets to polymerise, trapping the cells into polyacrylamide beads (Spencer et al., 2016). The method has been used for linking taxonomic information with the target dissimilatory sulphite reductase gene *dsrB* of a lake water microbial community (Spencer et al., 2016) as well as with antibiotic resistance associated genes in wastewater treatment plants (Hultman et al., 2018). Another non-microfluidics-based approach is to fix the bacterial cells as demonstrated by Kuchina et al. (2021) in microSPLiT (microbial split-pool ligation transcriptomics) and by Blattman et al. (2020) in PETRI-seq (Prokaryotic Expression-profiling by Tagging RNA In Situ and sequencing). In these methods, the mRNA (cDNA) of fixed and permeabilised bacterial cells is split-pool barcoded, resulting in the mRNA from each cell having a unique barcode. Split-pool barcoding is a method in which by repeated rounds of splitting, attaching a unique barcode, and pooling, each e.g. bacterial cell receives a unique combination of ligated barcodes. For example, if the barcoding is performed in a 96-well-plate-format three times, the number of potential unique barcodes is $96^3 = 884{,}736$.

1.2.3 Study of spatial organisation

Microbial communities exhibit specific spatial structures, which cannot be captured using methods that require homogenisation of the sample, such as metagenomics (Sheth et al., 2019). The spatial structure of a community can be observed directly using microscopy as in the HiPR-FISH (high-phylogenetic-resolution microbiome mapping by fluorescence in situ hybridization) method described by Shi et al. (2020), which utilises hundreds of 16S rRNA gene targeting fluorescently labelled probes. The method has been used to reveal the effect of antibiotic treatment on the spatial structuring of mouse gut microbiome as well as the temporal structural variation of human oral biofilm. The spatial associations of different bacteria can also be

observed using sequencing-based methods, such as MaPS-seq (metagenomic plot sampling by sequencing) by Sheth et al. (2019) and SAMPL-seq (Split-And-pool Metagenomic Plot-sampling sequencing) by Richardson et al. (2025), which have been used for the analysis of gut microbiomes in mice and humans, respectively. In both of the methods, the bacterial sample is embedded into a polymer matrix, which is then fractured and barcoded so that all of the particles have a unique barcode. While in MaPS-seq the barcoding is achieved using microfluidics and barcoded beads, in SAMPL-seq split-pool barcoding is used.

1.3 Aims of the thesis

The first aim of this thesis is to present how the common next-generation sequencing based approaches are currently often used for the analysis of bacterial communities, essentially to answer the question "who is in the community?". This question is answered in **Chapter I** and **Chapter II** by using the 16S rRNA gene amplicon sequencing.

In **Chapter I** the specific aim is to explore how endophytic bacterial communities respond to glyphosate-based herbicide (GBH) and phosphate treatments in various potato, faba bean and oat tissues in two different growth stages of the plant, the vegetative and the flowering stages. Additionally, based on the taxonomy, the potential sensitivity to glyphosate of the bacteria is determined using an *in silico*-based approach.

In **Chapter II**, the aim is to identify factors that affect the gut microbiota of two odonate species, the damselfly *Lestes sponsa* and the dragonfly *Sympetrum vulgatum*. Alongside four different environmental factors — host species, sex, habitat and season — the effect of diet, obtained using the sequencing of the cytochrome c oxidase subunit I (COI) gene, is investigated.

The second aim of this thesis is to develop and present novel tools and methods for the study of complex microbial communities, such as the ones studied in **Chapters I** and **II**, essentially to be able to answer the questions "who is doing what?" and "who is interacting with whom?". This is the shared aim of **Chapter III** and **Chapter IV**.

Chapter III presents Prider, a computational tool we recently developed, which enables efficient primer and probe designing for multiplex single-cell and screening applications with a variety of targets. The primary aim of Prider is to be able to efficiently design a minimal set of primer or probe sequences that have the highest possible coverage of the input DNA sequences, regardless of the number of target sequences in the input. Another key aim is ease of use with minimal user intervention required.

Chapter IV describes a single-cell sequencing method we recently developed, which combines porous polyacrylamide beads with split-pool barcoding for the analysis of both the microbial genetic heterogeneity in a community as well as the spatial structure. As there are already many methods capable of single cell analysis of microbiomes, one fundamental aim is to develop a method that is easy to set up and does not require specialised or expensive equipment or materials, such as those used in microfluidics.

2 Materials and Methods

Table 1 summarises the key methods used across this thesis and the chapters in which the methods have been utilised.

Table 1. List of methods used in different chapters of the thesis.

Method	Chapter
Plant tissue sampling	I
Insect sampling	II
DNA extraction	I, II
PCR	I, II, IV
Primer design	IV
Emulsion PCR	IV
Split-pool barcoding	IV
NGS read processing	I, II, IV
Alpha diversity analysis	I, II
Constrained analysis of principal coordinates (CAP)	I, II
Indicator species analysis	I, II
Differential abundance analysis	I, II
In silico EPSPS analysis	I
Benchmark testing	III

2.1 Chapter I

2.1.1 Fieldwork and data collection

The experimental field plots were located in southwestern Finland at the Ruissalo Botanical Garden (60°26′N, 22°10′E). The plots (23 m x 1.5 m each) consisted of 10 glyphosate-based herbicide (GBH)-treated plots and 10 water-treated control plots, which were divided in half: one half (subplot) treated with phosphate, the other with water, which resulted in four treatment groups: control, GBH, phosphate and GBH

with phosphate. The GBH-treated plots had been treated twice a year with GBH Roundup Gold for 7 years, the phosphate subplots with phosphate for 3 years, following the standard agricultural dosages as described by Helander et al. (2019). Briefly, in May 2020, prior to the experiment, each phosphate plot was sprayed with 80g of phosphate (Yara Ferticare) followed by each GBH plot being sprayed with 3 litres of Roundup Gold (450 g/isopropylamine glyphosate salt, CAS: 38641-94-0, 6.4 litres/ha in tap water) per plot. The control plots were sprayed with 3 litres of tap water. Two weeks after the treatments, 6 potato tubers (*Solanum tuberosum, var. 'Ditta'*), 15 oat seeds (*Avena sativa*) and 10 faba bean seeds (*Vicia faba*) were planted on each subplot.

Plant tissue samples of each treatment were collected from 10 replicate subplots: root and leaf samples from all the plants, nodules from the faba beans and tubers from the potatoes. The samples were collected at early (July 8th & 9th) and late summer (August 17th & 18th) during vegetative and end of flowering stage, respectively, except for potato tubers, which were only sampled in late summer. The plant samples were extracted using sterile scissors from the disease-free parts of the plant and kept in plastic bags on ice until they were washed with tap water and cut into 100 mg samples. Each sample was washed in 70% ethanol for a minute, 3% chloramine for 3 minutes and then rinsed three times in autoclaved MilliQ water for 1 minute. The samples were dried and stored at -80°C.

Frozen tissue samples were homogenised, and an Invisorb Spin Plant Mini Kit was used for the DNA extraction. The sample DNA concentrations were normalised for PCR. The V6–V8 region of the 16S rRNA gene was amplified, tagged and barcoded using nested PCR and amplicons sized 350–500 bp were selected for sequencing. The sequencing was conducted using the Ion 316™ Chip v2 in an Ion Personal Genome Machine (PGM™).

2.1.2 Data and statistical analyses

The 16S rRNA gene amplicon sequencing data was processed using Nextflow v.22.04.5 pipeline Ampliseq v.2.4.0 (Straub et al., 2020). Briefly, in the pipeline, FastQC (Andrews, 2010) was used for the raw read quality control, and the primer trimming was performed by Cutadapt (Martin, 2011). DADA2 (Callahan et al., 2016) and Barrnap (Seemann, 2018) were used to process the files into amplicon sequence variants (ASVs). The taxonomic classification was performed by DADA2 with the SILVA database (Yarza et al., 2014) as the reference.

The statistical analyses were performed using R v.4.4.3 (R Core Team, 2025) and Rstudio (Posit team, 2025) at ASV level unless otherwise stated and were conducted separately for each plant tissue and growth stage. The ASVs belonging to the domain Eukaryota or the phylum Cyanobacteria were filtered out. To identify if

the treatments had a significant effect on the alpha diversity of the plant endophytic communities, package mia v.1.14.0 (Ernst et al., 2022) was used for calculating the average Shannon diversity for each sample over 100 rarefaction rounds. The rarefaction depth was determined so that it was the same across different treatments within a plant tissue and growth stage and was the lowest possible read number that was at least 15% of the median read number. The significance of the differences between treatments was analysed using the two-sided Wilcoxon rank-sum test with holm multiple comparison correction using the packages ggpubr v.0.6.0 (Kassambara, 2023a) and rstatix v.0.7.2 (Kassambara, 2023b).

To identify if the treatments altered the bacterial community compositions, package vegan v.2.6-4 (Oksanen et al., 2022) was used to calculate the mean Bray-Curtis dissimilarity between treatments using the same rarefaction depths as with the alpha diversity analysis. Then, the constrained analysis of principal coordinates (CAP) was performed using the mean Bray-Curtis dissimilarity indices as the response variable and the GBH, phosphate and their interaction as the explanatory factors. The significance of the CAP results was tested using an ANOVA-like permutation test with 999 permutations.

Prior to any subsequent analysis, the samples were rarefied using the package mia v.1.14.0 using the same rarefaction depths as with the alpha diversity analysis. For the differential abundance analysis, only the ASVs with a prevalence of 0.1 were used and the packages ALDEx2 v.1.38.0 (Fernandes et al., 2014), ANCOMBC v.2.8.1 (Lin and Peddada, 2020), GuniFrac (ZicoSeq) v.1.8 (Yang and Chen, 2022), dacomp v.1.26 (Brill et al., 2022) and eBay v.0.1 (Liu et al., 2020) were used.

The potential sensitivity or resistance to glyphosate of the bacteria was evaluated using an EPSPS enzyme *in silico* analysis using the ATGC database (Kristensen et al., 2017) as described by Mathew et al. (2022). Cut-off scores of <0.2 and >0.8 (range 0–1) were used for resistant and sensitive, respectively. The significance of the differences in the abundance of the glyphosate-sensitive bacteria was evaluated using the Wilcoxon rank-sum test from the package rstatix v.0.7.2. The *p* values were corrected using the holm correction method.

2.2 Chapter II

2.2.1 Sample and data collection

In this chapter, two common target odonate species — *Lestes sponsa* (suborder Zygoptera) and *Sympetrum vulgatum* (Anisoptera) — were sampled in Järvelä (60°46′N, 22°39′E), Qvidja (60°30′N, 22°40′E) and Friskala (60°40′N, 22°23′E) in Southwest Finland. The sampling was conducted in 2020 during the odonate flight season at two time points: 26th–28th of July and 11th–13th of August. 18–22 adult

individuals of both species and sexes, in each location and time point were sampled, apart from in Friskala, where only *L. sponsa* was present. Once caught, each individual was placed in a sterile 10-ml collection tube with a piece of dampened paper towel for 24 hours. After the individual had defecated, the entire tube was frozen.

The DNA was extracted from each of the odonate faeces separately using a NucleoSpin Tissue XS Kit. To identify the odonate prey, two separate primer pairs were used to amplify the two halves of the cytochrome c oxidase subunit I (COI) gene. For the microbiota analysis, the V4 region of the bacterial 16S rRNA gene was amplified. The sequencing was conducted using Illumina MiSeq v3 PE 2×300 bp sequencing with two technical replicate sequencing runs.

2.2.2 Data and statistical analyses

Prior to the processing of the sequencing data, the prey COI gene read sets were combined and all subsequent steps were done separately for the prey and bacterial 16S rRNA gene sequencing data. The paired-end reads were merged and trimmed using VSEARCH v.2.14.2 (Rognes et al., 2016). The bacterial reads were truncated to 220 bp. Software cutadapt v.1.14 (Martin, 2011) was used for the primer removal and USEARCH v.11 (Edgar, 2010) for the removal of the singletons, denoising and chimera removal. The sets of zero-radius operational taxonomic units (ZOTUs) were mapped back to the original trimmed reads for the read counts using VSEARCH. ZOTUs from any sample that had fewer reads than the extraction or PCR controls were removed. The BOLD Systems database (Ratnasingham and Hebert, 2007) was used for the taxonomic classification of the prey (Vesterinen et al., 2020). The bacterial ZOTUs were classified using VSEARCH with the "16S RDP training set v18" USEARCH database. The reads assigned to the same taxa were collapsed and the taxa that were present in only one of the two technical replicates were filtered out. The read counts from the technical replicates were added together and any taxa with < 0.05% proportion of the total reads in a sample or sequence count < 2 were removed. Only the prey taxa assigned to phylum Arthropoda and bacterial taxa assigned to the domains Bacteria or Archaea were kept. The data was normalised using total sum scaling.

The statistical analyses of the data were conducted using R v.4.4.3 (R Core Team, 2025) and Rstudio (Posit team, 2025) at ZOTU level unless otherwise stated. The ASVs belonging to the class Cyanobacteria were filtered out. Dufrene-Legendre Indicator Species Analysis for both the prey and the bacterial data at genus-level was performed using R package labdsv v.2.1-0 (Roberts, 2023) and the *p* values were adjusted using the Benjamini-Hochberg correction. The package vegan v.2.6-4 (Oksanen et al., 2022) was used for calculating the prey (genus-level) and bacterial

Shannon diversities. The significance of the differences between the environmental factors — odonate species, sex, sampling site and sampling season — was analysed using the two-sided Wilcoxon rank-sum test with holm multiple comparison correction from the packages ggpubr v.0.6.0 (Kassambara, 2023a) and rstatix v.0.7.2 (Kassambara, 2023b). The Spearman rank correlation test was used to assess the correlation between the dietary and gut bacterial Shannon diversities.

To assess whether the environmental factors had an effect on the prey (species-level) or the gut microbiota, the constrained analysis of principal coordinates (CAP) was performed on Bray-Curtis dissimilarity indices using the package vegan v.2.6-4. For the prey, if no species-level information was available, ZOTU was used. As the site Friskala lacked *Lestes sponsa* samples, it was excluded from the analysis. The significance of the model and the individual factors was tested using an ANOVA-like permutation test with 999 permutations. As the sampling site and season did not meet the test assumptions, the significance of individual factors was only tested on odonate species and sex.

To identify if the prey had a significant effect on the gut microbiota assemblage and to identify some key prey taxa, vegan package method "ordistep" was used to build a model for CAP using the bacterial Bray-Curtis dissimilarity indices as the response variable and the prey abundances at the genus-level as the explanatory factors. Separate models were also built for *L. sponsa* and *S. vulgatum*. To assess the significance, an ANOVA-like permutation test with 999 permutations was used.

2.3 Chapter III

In this chapter a novel R package Prider was developed for the efficient designing of primers and probes from complex and heterogeneous DNA sequence sets for multiplexed assays. Prider was developed using C++11 and R v.4.0.5 (R Core Team, 2025) package Rcpp v.1.0.7 (Eddelbuettel and François, 2011). The processing speed of Prider was assessed using two randomly generated FASTA file sets, one containing 310 files with increasing number of bases per file (300 sequences each) and the other containing 300 files with increasing number of sequences per file (465,000 bases each). Each file contained 10 replicates of each number of bases or sequences, respectively. Prider is available at <https://github.com/tamminenlab/prider>.

2.4 Chapter IV

In this chapter a novel polyacrylamide bead-based laboratory method was developed and tested with the aim of being able to study complex microbial communities. The exact protocol is detailed in **Chapter IV**.

2.4.1 Polyacrylamide beads

Porous polyacrylamide beads were formed by combining 10 µL of bacterial cells (in 1xPBS), 52 µL 1xTE buffer, 5 µL of each of the three 100 µM acrydited forward primers, 3 µL 10% ammonium persulfate and 20 µL 30% acrylamide/bisacrylamide suspension (37,5/1) with a RAN-TEMED mixture containing 0.5 µL of >99% tetramethyl ethylenediamine in 100 µL of HFE7500 + 20g of 5 weight-% 008-FluoroSurfactant RAN oil. The mixture was emulsified with a pipette and incubated for approx. 17 hrs at 65°C. The cells used in the experiment were provided by Microbial Domain Biological Resource Center HAMBI (Biodiversity Collections Research Infrastructure (HUBCRI), Helsinki Institute of Life Science (HiLIFE)). One set of beads contained the cells belonging to the species *Comamonas testosteroni* and *Brevundimonas bullata* to mimic a physical interaction between the two species and the other set of beads contained only the species *Agrobacterium tumefaciens* to mimic the lack of interaction. After the incubation, the emulsion was broken down using 97% perfluoro-1-octanol, the beads were washed with 1xTE buffer and the two bead sets were combined. The beads were size filtered with a 70 µm cell strainer.

The V4 region of the 16S rRNA gene, a 299 bp long target specific to *A. tumefaciens* and a 283 bp long target specific to *C. testosteroni* of the cells encapsulated in polyacrylamide beads were amplified in a 20-cycle multiplex emulsion PCR in 16 replicates, with the amplicons attaching to the beads via the acrydited forward primer. The emulsion was based on 4% ABIL EM 90 and 0.05% Triton X-100 in mineral oil. After the PCR, the replicates were pooled, and the emulsion was broken using water-saturated diethyl ether and water-saturated diethyl acetate. The amplicons in the beads were made single stranded using NaOH and the beads were washed with 1xTE buffer.

2.4.2 Split-pool barcoding and sequencing

In split-pool barcoding, barcode oligonucleotides were attached to the amplicons created during the emulsion PCR via linker oligonucleotides. Briefly, in each of the three rounds of split-pool barcoding the polyacrylamide beads were split into a 96-well plate, each well containing a linker oligonucleotide and a unique barcode oligonucleotide. Prior to the barcoding, the linker and the barcode were hybridised so that the 5’ end of the linker was available for hybridisation with the amplicon and 3’ end of the barcode was available for the next linker in the subsequent round of barcoding. As the beads and the linker-barcode molecules were incubated, the linker hybridised with the amplicon in such a manner that the 3’ of the amplicon was adjacent to the 5’ end of the barcode and could be ligated using T4 DNA ligase. Subsequently, the ligase was inactivated with a heat-treatment and the linker-barcode

molecules that were not hybridised were blocked by hybridising a blocker oligonucleotide on the linker. After the barcoding, the beads were pooled and washed with 0.1% SDS and 1xTE buffer. The protocol was then repeated twice more to create a set of uniquely barcoded amplicons with 884,736 possible combinations of barcodes. Alongside containing the barcode sequence, the first set of barcode oligonucleotides contained a 10-nucleotide unique molecular identifier (UMI) sequence.

After the split-pool barcoding, the DNA was made single stranded with NaOH, and the beads were washed with 1xTE buffer. The DNA was detached from the polyacrylamide beads using USER enzyme, which cleaves the uracil nucleotide in the acrydited forward primer and was separated from the beads using a centrifuge. The barcoded DNA was amplified using two rounds of PCR and indexed for sequencing. The PCRs were conducted separately for the barcoded 16S rRNA and *A. tumefaciens*-specific target. The sequencing was conducted using Illumina MiSeq v2 PE 2x250 bp sequencing.

2.4.3 Data analysis

The adapter sequences were trimmed automatically by the Illumina FASTQ file generation pipeline. The quality of the raw reads was then evaluated using FastQC v0.12.1. (Andrews, 2010). Altogether, the sequencing returned 739,421 barcode-end reads, which were analysed using R v.4.4.3 (R Core Team, 2025) and RStudio (Posit team, 2025). Only the sequences with the expected barcode structure were included in the analysis. The sequences outside of the barcode sequence were mapped against the 16S rRNA gene sequences of *Agrobacterium tumefaciens*, *Brevundimonas bullata* and *Comamonas testosteroni* and the expected amplicon sequence of the *A. tumefaciens*-specific target. The annotated reads were filtered based on the 10-nt unique molecular identifier (UMI) sequence included in the first barcode oligonucleotide, so that only reads with a UMI that was clearly associated with just a single barcode and target were kept.

3 Results and discussion

3.1 Plant microbiome responses to agrochemicals

In **Chapter I**, the conventional method of amplification and sequencing of the hypervariable regions V6–V8 of the 16S rRNA gene was used. Using this approach, various changes in plant endophytic bacterial diversity and community composition in response to glyphosate-based herbicide (GBH) and phosphate as well as their combination (GP) were identified. The observed changes were not only limited to early or late growth stage of the plant or to one plant tissue but varied between these factors.

Phosphate-containing treatments reduced the bacterial alpha diversity (Shannon diversity) in the late growth stage potato leaves and tubers. Additionally, in the early growth stage oat leaves, both GBH and phosphate reduced the alpha diversity compared to the control. Some significant changes in response to the treatments in the endophytic bacterial community compositions were observed in the majority of the plant tissues and growth stages surveyed (Table 2). The shifts were solely driven by GBH in potato roots, where the effect was stronger in the early growth stage than in the late growth stage. The changes driven solely by phosphate were observed in the late growth stage potato leaves and tubers as well as late growth stage roots of oats. Both the GBH and phosphate treatments had a significant effect on the late growth stage faba bean roots. In the early growth stage faba bean leaves, phosphate was the major driver of the shifts, but some of the changes were specifically driven by the GP treatment. Finally, some shifts in the bacterial community composition were specifically driven by the GP treatment in the late growth stage oat leaves. In the majority of the cases, the treatments explained only a small amount of the variation observed. The shifts in the diversity and the microbiota were somewhat reflected in the abundance of the glyphosate-sensitive bacteria in some of the plant parts and growth stages. While the abundance of glyphosate-sensitive bacteria increased in the late growth stage potato leaves in response to phosphate and GP treatments, the abundance reduced in the early faba bean leaves in response to the phosphate treatment. Additionally, in the early growth stage oat leaves, GBH treatment reduced the abundance of sensitive bacteria.

Table 2. The variance in microbiota explained by the two first axes of the constrained analysis of principal coordinates and the significance level of glyphosate-based treatment (GBH), phosphate (P) and their combination/interaction (GP) for each plant part and growth stage where treatment had a significant effect on the bacterial community. Significance levels: *** $p \leq 0.001$; ** $p < 0.01$; * $p < 0.05$.

Plant / Part / Stage	Variance explained		Significance		
	CAP1	CAP2	GBH	P	GP
Potato Root Early	4.99 %	3.05 %	**		
Potato Root Late	3.78 %	2.01 %	*		
Faba bean Root Late	6.79 %	2.84 %	*	*	
Potato Leaf Late	19.06 %	1.84 %		***	
Potato Tuber	8.95 %	2.34 %		***	
Oat Root Late	5.87 %	2.55 %		*	
Faba bean Leaf Early	6.35 %	2.99 %		**	*
Oat Leaf Early	3.96 %	2.68 %			*

Previously it has been shown in lupine, strawberry and potato that the endophytic bacterial communities respond differently to the GBH and phosphate treatments in different plant tissues, and the results of **Chapter I** are in line with those observations (Ramula et al., 2022; Mathew et al., 2023, 2024). Additionally, shifts in the abundance of glyphosate-sensitive bacteria have previously been observed in strawberry and potato (Mathew et al., 2023, 2024). However, in the study conducted by Mathew et al. (2024) in the same site a year prior, GBH significantly affected the microbiota of the potato leaves, rather than roots, and the effect of phosphate was only seen in tubers, rather than leaves and tubers. Also, the shifts in the abundance of glyphosate-sensitive bacteria in response to the treatments differed. The conflicting results between the two studies could be caused by natural variations in the plant endophytic or soil microbial communities as well as differences in various environmental factors, such as the soil pH or the plant metabolites (Petrushin et al., 2024).

All the significant and major shifts in the abundance of individual bacterial taxa were in response to the phosphate treatments (Figure 1). Both of the phosphate treatments increased the abundance of *Sphingomonas faeni* and reduced the abundance of *Stenotrophomonas maltophilia* in the late growth stage potato leaves. In the late faba bean leaves, the effect of phosphate on the *S. maltophilia* was the opposite. Additionally, the phosphate treatment increased the abundance of *Allorhizobium-Neorhizobium-Pararhizobium-Rhizobium rhizogenes*. Bacteria of the

genus *Stenotrophomonas* are known for their beneficial health and growth-related interactions with the host plant as well as phosphate solubilisation (Ryan et al., 2009; Suliasih and Widawati, 2021). Bacteria belonging to the genus *Allorhizobium-Neorhizobium-Pararhizobium-Rhizobium* form symbiotic relationships in root nodules of leguminous plants and have previously been observed to become more abundant in the rhizosphere in response to an increase in available phosphorus (Guo et al., 2024).

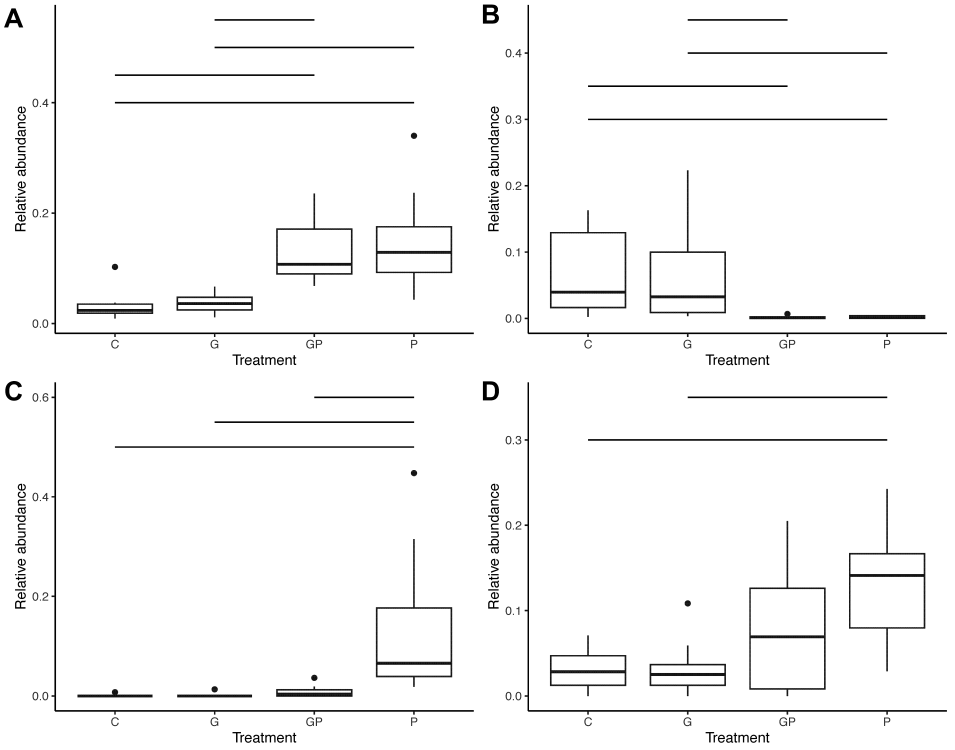

Figure 1. Relative abundances of (A) *Sphingomonas faeni* and (B) *Stenotrophomonas maltophilia* in late potato leaves, (C) *Allorhizobium-Neorhizobium-Pararhizobium-Rhizobium rhizogenes* in early and (D) *S. maltophilia* in late faba bean leaves. Comparing control (C), glyphosate-based herbicide (G), phosphate (P) and G with P (GP) treatments. Significant differences in abundance are indicated with a line and derive from multiple differential abundance analysis estimators. Figure adapted from **Chapter I**.

As many of the previous studies on the plant endophytic microbiomes have not considered the growth stage of the plant, one aim of this study was to reveal potential growth-stage-dependent differences in the microbiome responses. Although there were some exceptions, the majority of the changes driven by phosphate in potato and oats were observed in the late growth stage and the effects

of GBH were stronger in the early growth stage. Some of these effects could be explained by the translocation of phosphate, glyphosate residues and degradation products from the soil to the roots and to the leaves (Poirier and Bucher, 2002; Singh et al., 2020). Glyphosate can affect the microbiome either directly, by inhibiting the shikimate pathway or indirectly by changing the plant phenotype (Ruuskanen et al., 2023). The lack of monitoring of external factors, such as glyphosate residue and glyphosate degradation product (such as AMPA) levels, and internal factors, such as phytohormone levels, were major limitations of **Chapter I**. However, previous monitoring has revealed that glyphosate residues and AMPA are higher in the GBH-treated soil compared to the control soil 2–3 months after the pesticide application (Mathew et al., 2023). Although the mechanisms of action and wider consequences remain elusive, **Chapter I** showed that endophytic microbiomes of different plant tissues are affected by field-realistic phosphate and glyphosate-based herbicide treatments and that the effect is dependent on the growth stage of the plant.

3.2 Odonate gut microbiota, environment and diet

In **Chapter II**, the focus was on the gut microbiota of two carnivorous odonate species, the dragonfly *Sympetrum vulgatum* and the damselfly *Lestes sponsa*. The approach for determining the bacterial community assemblage was similar to the one used in **Chapter I**, but here the target for amplification and sequencing was the hypervariable region V4 of the 16S rRNA gene, as opposed to the V6–V8. Additionally, the recent diet (prey) of the odonates was determined by the amplification and sequencing of the cytochrome c oxidase subunit I (COI) gene. Both the diet and the gut microbiota were determined from faecal samples. The odonate species, sex (female/male), the sampling season (early/late) and site (Friskala/Qvidja/Järvelä), hereafter the environmental factors, and the diet of the odonates had a minor effect on the gut microbiota.

The majority of the odonate diet consisted of the order Diptera, which has also been previously observed in other studies (Kaunisto et al., 2017, 2020). The most common bacterial order found was Enterobacteriales. However, there was variation between the biological replicates in regard to both the diet and gut microbiota. Only one bacterial family, the family *Enterobacteriaceae*, had at least 1% abundance across all the samples. This family includes various genera and species known to be associated with the gut microbiotas of insects (Moran et al., 2005; Behar et al., 2005; Rizzi et al., 2013). Some prey and bacterial taxa had weak but significant associations with certain environmental factors. The aphid genus *Anoecia* was associated with *S. vulgatum* and the diptera *Microchironomus* with the early season odonates. Other more specific associations consisted of the ant genus *Lasius* being

associated with the early flying season *S. vulgatum* in Friskala, the Diptera *Bryophaenocladius* with the *S. vulgatum* in early Qvidja and *Sepsis* with the female *S. vulgatum* in early Friskala. These kinds of associations are potentially related to the availability of the food during the time of sampling. Additionally, the bacterial genus *Asaia* was associated with the site Friskala and genus *Ferruginibacter* with the *S. vulgatum* in the early season Friskala. The lack of strong associations and somewhat high variability between the biological replicates was also reflected in how much of the variation in the odonate diet or the gut microbiota the environmental factors explained: only 4% of the variation in the diet and 3.2% in the gut microbiota. Of the individual environmental factors, the odonate species was the only one to have a statistically significant, albeit small, effect on either. These results were somewhat in line with other dragonfly and damselfly gut microbiota studies where the odonate species or the sampling site did not have a strong effect on the host gut microbiota assemblage (Nobles and Jackson, 2020; Morrill et al., 2023).

The dietary and bacterial alpha (Shannon) diversities varied significantly between the odonates, sampling sites and sampling seasons, and in some cases followed the same patterns: both the dietary and bacterial diversities were higher in *S. vulgatum* than in *L. sponsa*, early Järvelä had the lowest diversities compared to the other sites as well as compared to the late season, and in Friskala the diversities were higher in the early season than in the late season. The actual correlation between the dietary and bacterial diversity was weak yet statistically significant: 11% for the whole data ($p = 0.025$) and 25% specifically in the early flight season ($p < 0.001$). This kind of correlation has previously been observed in other studies focusing on other species of odonates (Deb et al., 2019). The potential impact of diet on the gut microbiota was not limited to alpha diversity, but the abundance of various prey genera was also found to have a significant effect on the bacterial community compositions. Overall, the significant prey genera explained 11% of the variation in the gut microbiota, the four most influential genera being *Microchironomus*, *Hybos*, *Paratanytarsus* and *Endochironomus*. When the analysis was conducted separately for the two odonate species, the diet significantly explained 9.1% of the gut microbiota variation in *S. vulgatum* and 14.6% in *L. sponsa*. The diet had a stronger effect on the microbiota than the environmental factors. Similar results have previously been observed in *Drosophila melanogaster* (Chandler et al., 2011). Diet is considered as one of the potential horizontal transmission routes of gut microbiota (Lange et al., 2023) and although the proportion of variation in the gut microbiota explainable by the diet was somewhat low in this study, some level of horizontal transmission may have been observed. Insects are known for their intraspecific variation of gut microbiota (Lange et al., 2023), hence the variation in the prey microbiota may have resulted in the lack of any strong patterns between the odonate gut microbiota and the diet.

3.3 Methods for bacterial single-cell analysis

3.3.1 Prider: multiplexed primer and probe designing

In **Chapter III**, the aim was to develop and describe a novel R package called Prider (available at <https://github.com/tamminenlab/prider>), which addresses the need for highly scalable primer and probe design tools that are suitable for scripting and require minimal user intervention, particularly in the context of studying microbial communities using multiplex molecular methods. More specifically, the aim was to have a computationally efficient solution for the set cover problem (Shyu and Lee, 1990). This challenge occurs when designing primers or probes for multiplex molecular experiments, as it requires creating a minimal set of primers or probes to cover all sequences within a very large and heterogeneous set of target DNA, such as potential antimicrobial genes.

Prider was tested with a set of randomly generated input FASTA files. The coverage of primers designed by Prider was nearly optimal and the processing speed was constant and linearly dependent on the total number of bases in the input files, up to 9.03×10^6 bases. Scaling linearly to increasing data is especially useful in those fields of microbiology in which the number of targets of interest can be in the hundreds, such as surveillance of antibiotic resistance (Tamminen et al., 2020).

One key limitation of Prider is that the process of successful primer and probe designing is not only reliant on finding potential target sequences. Some of the primer/probe candidates designed may not be suitable for their intended application for various reasons, such as due to formation of secondary structures or inappropriate melting temperatures (Zangenberg et al., 1999). Furthermore, the package does not evaluate the potential amplicon lengths. Hence, although Prider requires minimal user input, the downstream steps of the primer or probe design process may require more effort. In practice Prider could for instance be used for an efficient production of large number of primer candidates, for example for a large set of known antimicrobial resistance genes, which are subsequently screened for their hybridization properties and amplicon lengths.

3.3.2 Polyacrylamide bead split-pool barcoding

The aim of **Chapter IV** was to develop and describe a novel method, polyacrylamide bead split-pool barcoding, that could be utilised for single-cell analysis of complex microbial communities for both the spatial analysis as well as linking genetic traits to single cells. The method described in **Chapter IV** combines various elements described previously in literature. Split-pool barcoding of eukaryotic single-cell transcriptomes was first described by Cao et al. (2017) and Rosenberg et al. (2018)

and since then, the method has been adapted for microbial single-cell RNA sequencing by Kuchina et al. (2021). Split-pool barcoding has also been previously combined with polyacrylamide beads for the barcoding of eukaryotic RNA in a microfluidic-based method, where the polyacrylamide beads are first barcoded without the presence of cells, and the cellular RNA or DNA is then barcoded using the beads (Delley and Abate, 2021). Here, the intention was to use split-pool barcoding to trace the origin of 16S rRNA gene amplicons and other PCR-amplifiable genomic loci to their cell of origin. The identification of appropriate loci and the conserved sequences within them for the PCR was facilitated by the Prider software package, developed in **Chapter III**.

An overview of the polyacrylamide bead split-pool method is summarised in Figure 2. *Brevundimonas bullata* and *Comamonas testosteroni* cells were encapsulated together to mimic a spatial nearness in a community, while *Agrobacterium tumefaciens* cells were encapsulated separately. In the multiplex emulsion PCR, prior to the split-pool barcoding, three barcoding targets were amplified: 16S rRNA gene region V4, a 299 bp long genomic target specific to *A.*

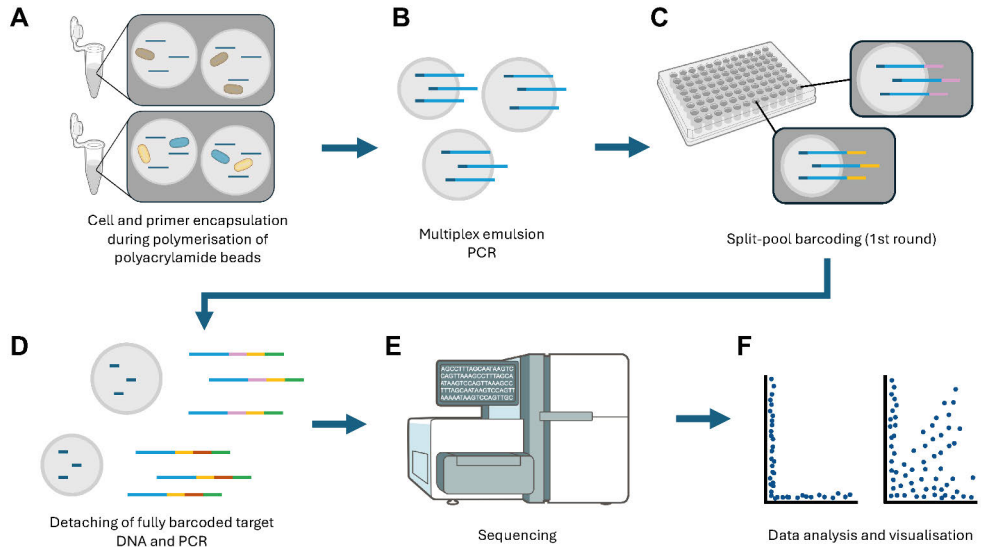

Figure 2. Polyacrylamide split-pool barcoding method overview. (A) Bacterial cells and acrydited primers were encapsulated in polyacrylamide beads by polymerising acrylamide-TEMED solution in emulsion droplets. In the present experiment, one set of beads (top) contained just one strain, the other set (bottom) contained two strains. (B) The target DNA was amplified in multiplex emulsion PCR, the amplicons attaching to the beads via the acrydited primers. (C) The amplicons were split-pool barcoded. (D) The barcoded DNA was detached from the beads and amplified and (E) sequenced with the Illumina MiSeq™ system. (F) The reads from the sequencing were analysed and barcode distributions visualised. The expected results for non-associated strains or targets on the left and associated strains or targets on the right. Figure created using illustrations from NIAID NIH BIOART. Figure adapted from **Chapter IV.**

tumefaciens and a 283 bp long genomic target specific to *C. testosteroni*, the last one not included in the post-split-pool barcoding PCRs due to issues with the amplicon.

3.3.2.1 Evaluation of the method

The structural integrity of the polyacrylamide beads and the successful encapsulation of the bacterial cells was evaluated microscopically. The lysis and the release of genetic material of the bacterial cells, which took place during the polyacrylamide bead polymerisation, was evaluated indirectly using gel electrophoresis, i.e. by the presence of correct sized bands, and confirmed with the sequencing data. The three bacterial strains used — *A. tumefaciens*, *B. bullata* and *C. testosteroni* — all are Gram-negative bacteria and therefore the lysis conditions were suitable for bacteria with a cell wall. However, as the study did not include for example any Gram-positive bacteria, biofilm structures or bacteria outside the phylum Pseudomonadota, it is difficult to evaluate the suitability of the method for all types of microbial communities, including those with certain spatial structures. Hence, an additional enzymatic lysis, such as the one described by Spencer et al. (2016) in the epicPCR protocol may be required.

The correct formation of the barcode was evaluated separately for the 16S rRNA gene and the *A. tumefaciens*-specific target reads. The correct structure of the barcode was defined as follows: it contained the expected non-variable regions between the barcode 3 (BC3) and BC2 variable regions, BC2 and BC1 variable regions and the BC1 variable region and the target amplicon, the variable regions were one of the 96 possible barcode sequences and the unique molecular identifier (UMI) situated in the first barcode oligo was of the correct length (Figure 3). The majority, 174,307 out of 242,826 barcodes in the 16S rRNA gene reads were correctly formed, while only 80,924 out of 496,595 barcodes in the *A. tumefaciens*-specific target reads were fully correct. The low amount of correctly formed product in the *A. tumefaciens*-specific target reads was mainly due to the lack of overhang region supposed to be situated between the first barcode oligo and the target amplicon. In those cases, the majority of the *A. tumefaciens*-specific target was also missing, which was reflected in the gel electrophoresis results in which the majority of the PCR product was approximately 300 bp shorter than expected. This issue could be explained by secondary binding sites of linker oligos or primers (Zangenberg et al., 1999), although no credible secondary sites were identified in the computational analysis of the oligos, or by secondary structures during the PCRs, potentially due to the high GC-richness (Frey et al., 2008) of the *A. tumefaciens*-specific target.

Figure 3. The structure of the full split-pool barcode attached to the target DNA. Different parts are indicated with different colours (from left to right): target DNA attached to the bead in dark blue, overhang region in blue, barcode 1 (BC1) in purple, BC2 in orange and BC3 in red. The 10-nt UMI region is indicated with vertical lines and the 8-nt variable regions in each BC with diagonal lines. Curly brackets indicate the non-variable regions used for the filtering of the barcodes. Figure from **Chapter IV**.

One key aim of the method described in **Chapter IV** was to be able to estimate the physical proximity of different bacterial species in a sample. This aspect of the method was tested by including only one type of bacteria (*A. tumefaciens*) in one set of the polyacrylamide beads and two types of bacteria (*C. testosteroni* and *B. bullata)* in the other set. In the event of a physical co-occurrence, the 16S rRNA gene sequences corresponding to the different cell types are expected to co-occur associated with a single barcode (Figure 2F, right). Conversely, for non-co-occurring cells, no other 16S rRNA gene sequences are expected to be associated with a single barcode (Figure 2F, left). For the analysis of the barcode-target associations, the three 8-nt-long variable regions from the BC1, BC2 and BC3 (Figure 3) were concatenated to form one 24-nt barcode sequence, resulting in 1,583 unique sequences. The sequencing data showed that *C. testosteroni* and *B. bullata* shared 184 unique barcodes, while *A. tumefaciens* was associated with *C. testosteroni* 47 and with *B. bullata* 31 times (Figure 4). To put these values in perspective, 262 unique barcodes were exclusively associated with *A. tumefaciens*, 258 with *B. bullata*, 538 with *C. testosteroni* and 118 with *A. tumefaciens*-specific genomic target. The second key aim of the method was to be able to link a genomic target, in this case the *A. tumefaciens*-specific genomic target, with the taxonomic information obtained from the 16S rRNA gene. For a given genetic target occurring within a particular cell type, it is expected to observe the sequence corresponding to this genetic target in a barcode shared with the corresponding 16S rRNA gene sequence (Figure 2F, right). The *A. tumefaciens*-specific target shared a unique barcode with *A. tumefaciens* 51 times, whereas the number of shared barcodes were 2 and 12 with *B. bullata* and *C. testosteroni*, respectively (Figure 4). Altogether, these results indicate that the phylogenetic localization of a genetic target can be traced back to the relevant cell type, with the noise caused by the crosstalk between the beads remaining minimal albeit not non-existent.

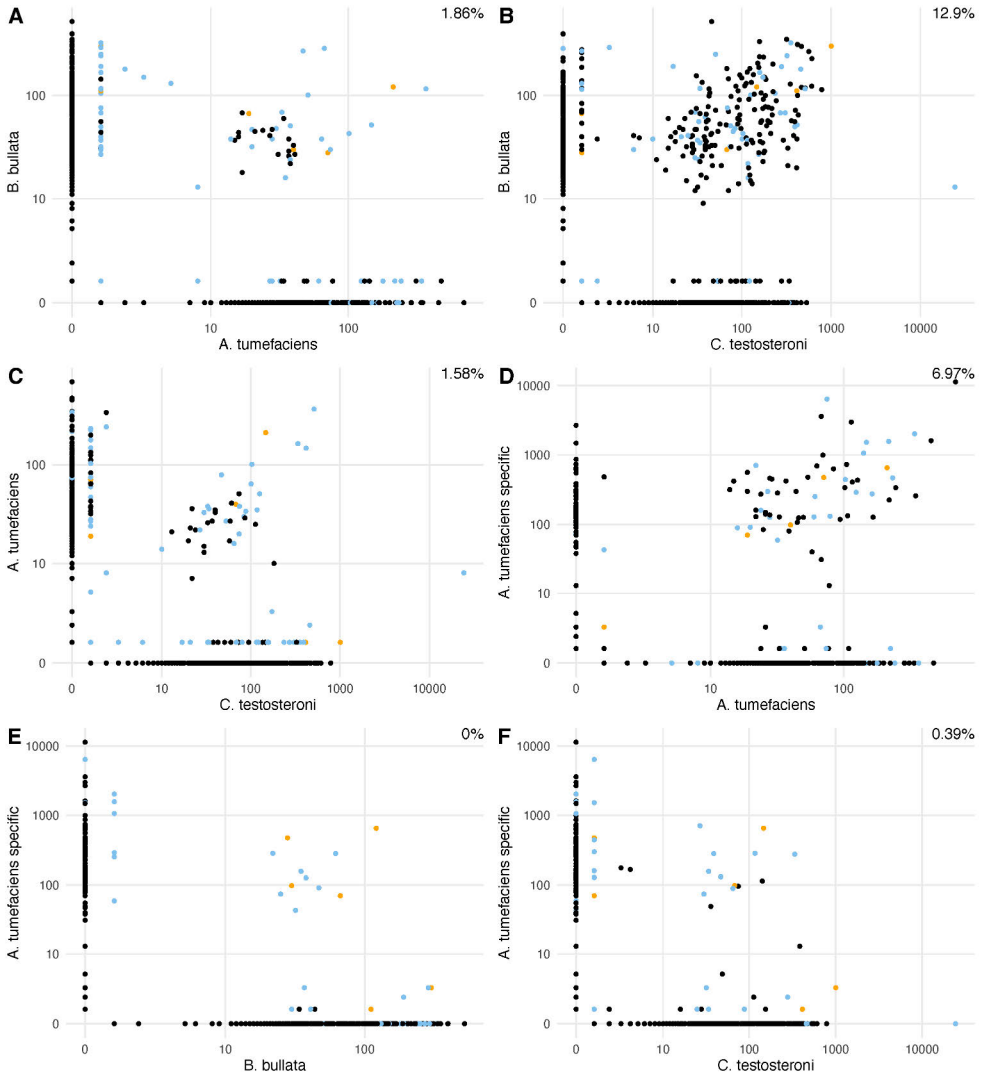

Figure 4. Barcode distributions and associations with the targets. Each point represents a barcode and its position the number of times it was associated with (A) *A. tumefaciens* or *B. bullata*, (B) *C. testosteroni* or *B. bullata*, (C) *C. testosteroni* or *A. tumefaciens*, (D) *A. tumefaciens* or *A. tumefaciens*-specific target, (E) *B. bullata* or *A. tumefaciens*-specific target and (F) *C. testosteroni* or *A. tumefaciens*-specific target. Panels B and D are expected to have a higher proportion of points that are ≥ 10 for both targets, while the other panels ideally should have none. Blue points represent barcodes associated with any three targets (n = 74), orange points with all four targets (n = 6). Ideally the number of these points should be minimal. Percentages are the proportion of barcodes that were solely associated ≥ 10 times with both respective targets. X and Y axes are on pseudo-logarithmic scale. Figure adapted from **Chapter IV**.

3.3.2.2 Future perspectives

The aim of **Chapter IV** was to demonstrate a high-throughput polyacrylamide bead split-pool barcoding method for studying both the physical interactions as well as genetic heterogeneity in microbial communities. Although the method performed well with the system used in this study, three key aspects still need to be further optimised. Firstly, the polyacrylamide bead split-pool barcoding method was tested with a three-species mock bacterial community. As natural microbial communities can be extremely diverse (Leviatan et al., 2022; Labouyrie et al., 2023), the performance of the method with highly complex communities needs to be tested in the future, initially with a more diverse mock community. Secondly, during the development and optimisation of the method, an excess number of cells was used, resulting in multiple cells per polyacrylamide bead. To be able to capture intra-species genetic heterogeneity, the method needs to be optimised to a single-cell resolution, thus the average number of cells per bead needs to be less than one. It is yet to be established whether an appropriately low cell density can be determined microscopically or whether flow cytometry is required. Thirdly, the number of samples that can realistically be barcoded using the method is currently very limited, as both the emulsion PCR and the split-pool barcoding are laborious workflows, and each sample needs to be processed separately. Hence, an additional sample-specific tagging should be included in the process, for example as part of the polyacrylamide bead polymerisation step.

Other potential future aspects to consider include re-evaluation and expansion of the targets of interest. During the optimisation process, the V4 region of the 16S rRNA gene and genomic targets specific to the bacterial strains used were barcoded. However, the taxonomic resolution of the V4 region may not be sufficient enough for natural bacterial communities (Johnson et al., 2019), hence, including other hypervariable regions of the 16S rRNA gene may improve the accuracy. Expanding the number and scope of the targets of interest will also require further optimisation and will be highly dependent on the type of model system used in future experiments. Furthermore, as the polyacrylamide bead size was only controlled by using a 70 µm cell strainer, the remaining beads ranged significantly in size, the smaller ones being approximately 20 µm in diameter. Thus, it may be worth investigating how different bead sizes affect the results of the method, as larger beads may be able to hold more DNA than the smaller beads.

3.4 Improving microbiological research with novel methods

In **Chapter I**, the endophytic bacterial community compositions were determined using 16S rRNA gene amplicon sequencing and their sensitivity to glyphosate was

determined using an *in silico* analysis, based on the taxonomic conservation of EPSP synthase. However, as a single mutation in the EPSP synthase active site can change the status from sensitive to resistant and vice versa (Rainio et al., 2021), the *in silico* analysis may not be completely accurate for all of the bacteria. This is potentially supported by a study on bumblebee gut microbiome in which the abundance of the bacteria thought to be glyphosate-sensitive significantly increased in abundance in response to glyphosate treatment (Helander et al., 2023). Bulk sequencing methods, such as the 16S rRNA gene amplicon sequencing utilised in **Chapter I**, or metagenomic sequencing cannot capture the intraspecies heterogeneity of bacteria (Brehm-Stecher and Johnson, 2004) leaving the associations between different genetic variants and different bacteria ambiguous. To be able to completely capture the variation in the EPSP synthase and heterogeneity in the glyphosate-sensitivity of bacteria, single-cell methods, such as the polyacrylamide bead split-pool method described in **Chapter IV**, are required. This approach would first require designing primers that amplify an appropriate target or targets within the gene encoding the EPSP synthase, which could then be barcoded simultaneously with the 16S rRNA gene target. Using the polyacrylamide bead split-pooling, the research could also be expanded to other targets of interest such as bacterial genes and pathways that are related to the health of the plant, allowing investigation of agrochemical-induced shifts in the bacteria capable of protecting its host from infection or causing infection (Rahme et al., 1995; Pan et al., 2008). Additionally, as some bacteria are capable of degrading glyphosate, other targets of interest could be found within pathways involved in the glyphosate metabolism (Dick and Quinn, 1995). These could for instance help to explain the shifts in the endophytic microbiota. It may also be possible to extend the analysis to other common targets within the microbiome, such as fungi. The design of the new primers could be facilitated by the Prider package described in **Chapter III**.

In **Chapter II** the gut microbiota of two odonate species was determined using the 16S rRNA gene amplification and sequencing approach. No specific features of the microbiome were attempted to be inferred from the taxonomy in this study, like in **Chapter I**, but there are several known genetic potential targets of interest relevant to the digestion and health of insects in bacteria. Suitable additional targets could be within for example useful bacterial pathways related to the breakdown of complex carbohydrates, micronutrient synthesis or protection against pathogens (Saqib et al., 2023) or related to pathogenic bacterial activities such as virulence factors or genes related to toxin production (Castagnola and Stock, 2014). Expanding the research to these kinds of targets using the polyacrylamide split-pool barcoding method would provide a deeper understanding of the functional potential of the odonate microbiomes.

Bacteria are known to form specific spatial structures as a result of beneficial interactions and active avoidance between cells (Nielsen et al., 2000). Although this aspect was not the focal point of either **Chapter I** or **Chapter II**, investigating these bacterial associations using the polyacrylamide bead split-pool barcoding method could provide insights into the spatial biogeography and bacterial associations in both plant bacterial communities and insect gut microbiomes (Smith et al., 2017; Cao et al., 2023), enhancing the understanding of how microbial communities are shaped in a variety of environments. Other spatial ecology methods have previously observed changes in the bacterial associations in response to dietary alterations in human and mouse gut microbiomes, demonstrating how external factors can impact the organisation of microbial communities (Sheth et al., 2019; Richardson et al., 2025). Observing the spatial biogeography of endophytic communities and insect gut microbiomes would be especially interesting, as bacteria differ in their capability to utilise and break down different substances, including glyphosate (Manogaran et al., 2018; de Lima et al., 2024). The ability to break down nutrients, glyphosate or other substances may not only benefit the bacteria possessing these capabilities, but also the surrounding bacteria, by either removing harmful substances from their environment or by facilitating metabolic cross-feeding.

4 Conclusion

Bacteria exist almost everywhere and affect the well-being and health of nearly everything living on Earth (Cockell, 2021; Ma et al., 2023; Virgo et al., 2025). They form heterogeneous microbiomes, multispecies communities, with specific spatial structures where they interact with each other and their environment (Yanni et al., 2019; Andam, 2019; Berg et al., 2020). In my thesis I have showcased the information that can be extracted from these communities using conventional methods and what may be possible in the future with new methods. In **Chapter I**, I identified various growth stage-dependent shifts in the diversity and the assemblage of the endophytic microbiota of potato, faba bean and oat in response to glyphosate-based herbicide and phosphate. In **Chapter II**, I explored how different environmental factors and diet affect the odonate gut microbiota. This chapter provided evidence that although the gut microbiota varies greatly between individual odonates, both environmental factors and diet have a minor effect on both the diversity and the assemblage of the community. **Chapter III** and **Chapter IV** presented novel tools we have developed to aid and improve the analysis of microbial communities. **Chapter III** presented a computational primer and probe designing tool Prider, specifically aimed for the analysis of complex and heterogeneous microbiomes. **Chapter IV** presented a laboratory method, polyacrylamide bead split-pool barcoding, designed for capturing both the genetic heterogeneity and spatial structuring of microbial communities.

Microbiome research has most often relied on the amplification and sequencing of the 16S rRNA gene, answering the question “who is there?” (Bose and Moore, 2023; Bartoš et al., 2024). Although this approach has produced an invaluable amount of new information about microbial communities across a plethora of systems, including the ones described in **Chapter I** and **Chapter II,** a much better understanding of the microbiomes can be gained using high-throughput single-cell methods (Brehm-Stecher and Johnson, 2004; Tedersoo et al., 2021; Stromberg et al., 2023). Additionally, efficient computational tools are required for the designing stage of new experiments (Kreer et al., 2020). The use of these new methods allows us to ask and answer question such as “who is doing what?” and “who is interacting with whom?”, extending both the amount and the scope of insights we can extract from a variety of microbiomes.

Acknowledgements

First of all, I would like to express my gratitude to my supervisor, Manu Tamminen. The insights and knowledge I have gained from you throughout this journey as well as your optimism and ideas have been invaluable. I am particularly thankful for the time you dedicated to reviewing all the manuscripts and my thesis, as well as for introducing me to all the people who ended up being my co-authors.

I would also like to extend my sincere thanks to my research director Prof. Harri Savilahti and my second supervisor Prof. Teppo Hiltunen for taking your time to provide feedback on my thesis. This research was funded by the Research Council of Finland and Doctoral Programme in Biology, Geography and Geology (BGG), and I am grateful for their support.

A special acknowledgment goes to Prof. Ellen Decaestecker and Dr. Derek Lundberg for acting as the pre-examiners of my thesis. Thank you for your thorough evaluations. I also wish to recognise Assoc. Prof. Lars Behrendt for agreeing to be my opponent.

I am deeply appreciative to my co-authors, from whom I have learned so much about plant microbiomes and odonates. Your collaboration has not only enriched my understanding but also made me a better researcher. In particular I would like to acknowledge Helandersaikkonenlab, Suni Mathew and Benjamin Fuchs for the plant microbiome data as well as Kari Kaunisto and Eero Vesterinen for the odonate data.

Warmest thanks to all the people who have worked at the Tamminen lab over the years; in particular Santeri, Jasmin, Elsi and Aditya, it has been a joy to work with you. I would also like to express my gratitude to my advisory group for their encouraging words and insights. Additionally, I wish to acknowledge the Department of Biology at the University of Turku, where my journey to become a researcher started in 2015 when I started my bachelor's studies.

To my family and friends (and my cats), your support — be it financial, in childcare, or simply the right words at the right moment — has been crucial. I owe a special debt of gratitude to Paul for your unwavering support and for being such an amazing partner throughout this endeavour. My dear daughter, Rosa, thank you for being the light of my life, inspiring me every day with your joy and curiosity.

December 2025
Niina Smolander

List of References

Ackermann, M., 2015. A functional perspective on phenotypic heterogeneity in microorganisms. *Nat Rev Microbiol*, 13, 497–508. https://doi.org/10.1038/nrmicro3491.

Andam, C. P., 2019. Clonal yet Different: Understanding the Causes of Genomic Heterogeneity in Microbial Species and Impacts on Public Health. *mSystems*, 4. https://doi.org/10.1128/msystems.00097-19.

Andrews, S., 2010. FastQC: A Quality Control Tool for High Throughput Sequence Data. <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>

Arnault, G., Mony, C., Vandenkoornhuyse, P., 2023. Plant microbiota dysbiosis and the Anna Karenina Principle. *Trends Plant Sci*, 28, 18–30. https://doi.org/10.1016/J.TPLANTS.2022.08.012.

Bai, S., Yao, Z., Raza, M. F., Cai, Z., Zhang, H., 2021. Regulatory mechanisms of microbial homeostasis in insect gut. *Insect Sci*, 28, 286–301. https://doi.org/10.1111/1744-7917.12868.

Balaban, N. Q., Helaine, S., Lewis, K., Ackermann, M., Aldridge, B., Andersson, D. I., Brynildsen, M. P., Bumann, D., Camilli, A., Collins, J. J., Dehio, C., Fortune, S., Ghigo, J. M., Hardt, W. D., Harms, A., Heinemann, M., Hung, D. T., Jenal, U., Levin, B. R., Michiels, J., Storz, G., Tan, M. W., Tenson, T., Van Melderen, L., Zinkernagel, A., 2019. Definitions and guidelines for research on antibiotic persistence. *Nat Rev Microbiol*, 17, 441–448. https://doi.org/10.1038/s41579-019-0196-3.

Balloux, F., van Dorp, L., 2017. Q&A: What are pathogens, and what have they done to and for us? *BMC Biol*, 15, 91. https://doi.org/10.1186/s12915-017-0433-z.

Bartlett, A., Padfield, D., Lear, L., Bendall, R., Vos, M., 2022. A comprehensive list of bacterial pathogens infecting humans. *Microbiology*, 168, 001269. https://doi.org/10.1099/mic.0.001269.

Bartoš, O., Chmel, M., Swierczková, I., 2024. The overlooked evolutionary dynamics of 16S rRNA revises its role as the "gold standard" for bacterial species identification. *Sci Rep*, 14, 1–9. https://doi.org/10.1038/s41598-024-59667-3.

Behar, A., Yuval, B., Jurkevitch, E., 2005. Enterobacteria-mediated nitrogen fixation in natural populations of the fruit fly Ceratitis capitata. *Mol Ecol*, 14, 2637–2643. https://doi.org/10.1111/j.1365-294X.2005.02615.x.

Berg, G., Rybakova, D., Fischer, D., Cernava, T., Vergès, M. C. C., Charles, T., Chen, X., Cocolin, L., Eversole, K., Corral, G. H., Kazou, M., Kinkel, L., Lange, L., Lima, N., Loy, A., Macklin, J. A., Maguin, E., Mauchline, T., McClure, R., Mitter, B., Ryan, M., Sarand, I., Smidt, H., Schelkle, B., Roume, H., Kiran, G. S., Selvin, J., Souza, R. S. C. de, Van Overbeek, L., Singh, B. K., Wagner, M., Walsh, A., Sessitsch, A., Schloter, M., 2020. Microbiome definition re-visited: old concepts and new challenges. *Microbiome*, 8, 103. https://doi.org/10.1186/s40168-020-00875-0.

Blattman, S. B., Jiang, W., Oikonomou, P., Tavazoie, S., 2020. Prokaryotic Single-Cell RNA Sequencing by In Situ Combinatorial Indexing. *Nat Microbiol*, 5, 1201. https://doi.org/10.1038/S41564-020-0729-6.

Bose, N., Moore, S. D., 2023. Variable Region Sequences Influence 16S rRNA Performance. *Microbiol Spectr*, 11. https://doi.org/10.1128/spectrum.01252-23.

Bottery, M. J., Pitchford, J. W., Friman, V. P., 2021. Ecology and evolution of antimicrobial resistance in bacterial communities. *ISME J*, 15, 939–948. https://doi.org/10.1038/s41396-020-00832-7.

Brehm-Stecher, B. F., Johnson, E. A., 2004. Single-Cell Microbiology: Tools, Technologies, and Applications. *Microbiol Mol Biol Rev*, 68, 538–559. https://doi.org/10.1128/MMBR.68.3.538-559.2004.

Brill, B., Amir, A., Heller, R., 2022. Testing for differential abundance in compositional counts data, with application to microbiome studies. *Ann Appl Stat*, 16, 2648–2671. https://doi.org/10.1214/22-AOAS1607.

Buetas, E., Jordán-López, M., López-Roldán, A., D'Auria, G., Martínez-Priego, L., De Marco, G., Carda-Diéguez, M., Mira, A., 2024. Full-length 16S rRNA gene sequencing by PacBio improves taxonomic resolution in human microbiome samples. *BMC Genomics*, 25, 310. https://doi.org/10.1186/s12864-024-10213-5.

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

Cao, J., Packer, J. S., Ramani, V., Cusanovich, D. A., Huynh, C., Daza, R., Qiu, X., Lee, C., Furlan, S. N., Steemers, F. J., Adey, A., Waterston, R. H., Trapnell, C., Shendure, J., 2017. Comprehensive single-cell transcriptional profiling of a multicellular organism. *Science*, 357, 661–667. https://doi.org/10.1126/science.aam8940.

Cao, Z., Zuo, W., Wang, L., Chen, J., Qu, Z., Jin, F., Dai, L., 2023. Spatial profiling of microbial communities by sequential FISH with error-robust encoding. *Nat Commun*, 14, 1477. https://doi.org/10.1038/s41467-023-37188-3.

Carrión, V. J., Perez-Jaramillo, J., Cordovez, V., Tracanna, V., De Hollander, M., Ruiz-Buck, D., Mendes, L. W., van Ijcken, W. F. J., Gomez-Exposito, R., Elsayed, S. S., Mohanraju, P., Arifah, A., van der Oost, J., Paulson, J. N., Mendes, R., van Wezel, G. P., Medema, M. H., Raaijmakers, J. M., 2019. Pathogen-induced activation of disease-suppressive functions in the endophytic root microbiome. *Science*, 366, 606–612. https://doi.org/10.1126/science.aaw9285.

Castagnola, A., Stock, S. P., 2014. Common Virulence Factors and Tissue Targets of Entomopathogenic Bacteria for Biological Control of Lepidopteran Pests. *Insects*, 5, 166. https://doi.org/10.3390/INSECTS5010139.

Chandler, J. A., Lang, J., Bhatnagar, S., Eisen, J. A., Kopp, A., 2011. Bacterial Communities of Diverse Drosophila Species: Ecological Context of a Host–Microbe Model System. *PLoS Genet*, 7, e1002272. https://doi.org/10.1371/JOURNAL.PGEN.1002272.

Cockell, C. S., 2021. Are microorganisms everywhere they can be? *Environ Microbiol*, 23, 6355. https://doi.org/10.1111/1462-2920.15825.

Compant, S., Clément, C., Sessitsch, A., 2010. Plant growth-promoting bacteria in the rhizo- and endosphere of plants: Their role, colonization, mechanisms involved and prospects for utilization. *Soil Biol Biochem*, 42, 669–678. https://doi.org/10.1016/J.SOILBIO.2009.11.024.

Coon, K. L., Valzania, L., Brown, M. R., Strand, M. R., 2020. Predaceous Toxorhynchites mosquitoes require a living gut microbiota to develop. *Proc R Soc B: Biol Sci*, 287, 20192705. https://doi.org/10.1098/RSPB.2019.2705.

Costerton, J. W., Lewandowski, Z., Caldwell, D. E., Korber, D. R., Lappin-Scott, H. M., 1995. Microbial biofilms. *Annu Rev Microbiol*, 49, 711–745. https://doi.org/10.1146/ANNUREV.MI.49.100195.003431.

Croucher, N. J., Coupland, P. G., Stevenson, A. E., Callendrello, A., Bentley, S. D., Hanage, W. P., 2014. Diversification of bacterial genome content through distinct mechanisms over different timescales. *Nat Commun*, 5, 5471. https://doi.org/10.1038/ncomms6471.

Dakal, T. C., Xu, C., Kumar, A., 2025. Advanced computational tools, artificial intelligence and machine-learning approaches in gut microbiota and biomarker identification. *Front Med Technol*, 6, 1434799. https://doi.org/10.3389/fmedt.2024.1434799.

Dastogeer, K. M. G., Tumpa, F. H., Sultana, A., Akter, M. A., Chakraborty, A., 2020. Plant microbiome–an account of the factors that shape community composition and diversity. *Curr Plant Biol*, 23, 100161. https://doi.org/10.1016/J.CPB.2020.100161.

David, H. L., 1970. Probability Distribution of Drug-Resistant Mutants in Unselected Populations of Mycobacterium tuberculosis. *Appl Microbiol*, 20, 810–814. https://doi.org/10.1128/AM.20.5.810-814.1970.

Deb, R., Nair, A., Agashe, D., 2019. Host dietary specialization and neutral assembly shape gut bacterial communities of wild dragonflies. *PeerJ*, 7, e8058. https://doi.org/10.7717/PEERJ.8058.

Delley, C. L., Abate, A. R., 2021. Modular barcode beads for microfluidic single cell genomics. *Sci Rep*, 11, 1–9. https://doi.org/10.1038/s41598-021-90255-x.

Dick, R. E., Quinn, J. P., 1995. Glyphosate-degrading isolates from environmental samples: occurrence and pathways of degradation. *Appl Microbiol Biotechnol*, 43, 545–550. https://doi.org/10.1007/BF00218464.

Dosch, C., Manigk, A., Streicher, T., Tehel, A., Paxton, R. J., Tragust, S., 2021. The gut microbiota can provide viral tolerance in the honey bee. *Microorganisms*, 9, 871. https://doi.org/10.3390/microorganisms9040871.

Eddelbuettel, D., François, R., 2011. Rcpp: Seamless R and C++ Integration. *J Stat Softw*, 40 https://doi.org/10.18637/JSS.V040.I08.

Edgar, R. C., 2010. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics*, 26, 2460–2461. https://doi.org/10.1093/BIOINFORMATICS/BTQ461.

Elowitz, M. B., Levine, A. J., Siggia, E. D., Swain, P. S., 2002. Stochastic gene expression in a single cell. *Science*, 297, 1183–1186. https://doi.org/10.1126/science.1070919.

Elvira, K. S., Gielen, F., Tsai, S. S. H., Nightingale, A. M., 2022. Materials and methods for droplet microfluidic device fabrication. *Lab Chip*, 22, 875. https://doi.org/10.1039/D1LC00836F.

Ernst, F. G. M., Shetty, S. A., Borman, T., Lahti, L., 2022. mia: Microbiome analysis. <https://microbiome.github.io/mia/>

Faust, K., Raes, J., 2012. Microbial interactions: from networks to models. *Nat Rev Microbiol*, 10, 538–550. https://doi.org/10.1038/nrmicro2832.

Feng, H., Zhang, N., Fu, R., Liu, Y., Krell, T., Du, W., Shao, J., Shen, Q., Zhang, R., 2019. Recognition of dominant attractants by key chemoreceptors mediates recruitment of plant growth-promoting rhizobacteria. *Environ Microbiol*, 21, 402–415. https://doi.org/10.1111/1462-2920.14472.

Fernandes, D., A., Reid, J., Macklaim, M., J., McMurrough, T.A, Edgell, D.R., Gloor, B., G., 2014. Unifying the analysis of high-throughput sequencing datasets: characterizing RNA-seq, 16S rRNA gene sequencing and selective growth experiments by compositional data analysis. *Microbiome*, 2, 15.

Ferrarini, M. G., Dell'Aglio, E., Vallier, A., Balmand, S., Vincent-Monégat, C., Hughes, S., Gillet, B., Parisot, N., Zaidman-Rémy, A., Vieira, C., Heddi, A., Rebollo, R., 2022. Efficient compartmentalization in insect bacteriomes protects symbiotic bacteria from host immune system. *Microbiome*, 10, 156. https://doi.org/10.1186/s40168-022-01334-8.

France, M. T., Cornea, A., Kehlet-Delgado, H., Forney, L. J., 2018. Spatial structure facilitates the accumulation and persistence of antibiotic-resistant mutants in biofilms. *Evol Appl*, 12, 507. https://doi.org/10.1111/EVA.12728.

Frey, U. H., Bachmann, H. S., Peters, J., Siffert, W., 2008. PCR-amplification of GC-rich regions: 'Slowdown PCR'. *Nat Protoc*, 3, 1312–1317. https://doi.org/10.1038/nprot.2008.112.

Fu, Y., Ho, B. T., Mekalanos, J. J., 2018. Tracking Vibrio cholerae cell-cell interactions during infection reveals bacterial population dynamics within intestinal microenvironments. *Cell Host Microbe*, 23, 281. https://doi.org/10.1016/J.CHOM.2017.12.006.

Galloway-Peña, J., Hanson, B., 2020. Tools for Analysis of the Microbiome. *Dig Dis Sci*, 65, 674–685. https://doi.org/10.1007/s10620-020-06091-y.

Guo, T., Yao, X., Wu, K., Guo, A., Yao, Y., 2024. Response of the rhizosphere soil microbial diversity to different nitrogen and phosphorus application rates in a hulless barley and pea mixed-cropping system. *Appl Soil Ecol*, 195, 105262. https://doi.org/10.1016/J.APSOIL.2023.105262.

Hammer, T. J., Sanders, J. G., Fierer, N., 2019. Not all animals need a microbiome. *FEMS Microbiol Lett*, 366, 117. https://doi.org/10.1093/FEMSLE/FNZ117.

Helander, M., Pauna, A., Saikkonen, K., Saloniemi, I., 2019. Glyphosate residues in soil affect crop plant germination and growth. *Sci Rep*, 9, 19653. https://doi.org/10.1038/s41598-019-56195-3.

Helander, M., Jeevannavar, A., Kaakinen, K., Mathew, S. A., Saikkonen, K., Fuchs, B., Puigbò, P., Loukola, O. J., Tamminen, M., 2023. Glyphosate and a glyphosate-based herbicide affect bumblebee gut microbiota. *FEMS Microbiol Ecol*, 99, fiad065. https://doi.org/10.1093/FEMSEC/FIAD065.

Hillmann, B., Al-Ghalith, G. A., Shields-Cutler, R. R., Zhu, Q., Gohl, D. M., Beckman, K. B., Knight, R., Knights, D., 2018. Evaluating the Information Content of Shallow Shotgun Metagenomics. *mSystems*, 3, e00069-18. https://doi.org/10.1128/MSYSTEMS.00069-18.

Hultman, J., Tamminen, M., Pärnänen, K., Cairns, J., Karkman, A., Virta, M., 2018. Host range of antibiotic resistance genes in wastewater treatment plant influent and effluent. *FEMS Microbiol Ecol*, 94, fiy038. https://doi.org/10.1093/FEMSEC/FIY038.

Ide, K., Saeki, T., Arikawa, K., Yoda, T., Endoh, T., Matsuhashi, A., Takeyama, H., Hosokawa, M., 2022. Exploring strain diversity of dominant human skin bacterial species using single-cell genome sequencing. *Front Microbiol*, 13, 955404. https://doi.org/10.3389/fmicb.2022.955404.

Iwai, S., Weinmaier, T., Schmidt, B. L., Albertson, D. G., Poloso, N. J., Dabbagh, K., DeSantis, T. Z., 2016. Piphillin: Improved Prediction of Metagenomic Content by Direct Inference from Human Microbiomes. *PLoS One*, 11, e0166104. https://doi.org/10.1371/JOURNAL.PONE.0166104.

Johnson, J. S., Spakowicz, D. J., Hong, B. Y., Petersen, L. M., Demkowicz, P., Chen, L., Leopold, S. R., Hanson, B. M., Agresta, H. O., Gerstein, M., Sodergren, E., Weinstock, G. M., 2019. Evaluation of 16S rRNA gene sequencing for species and strain-level microbiome analysis. *Nat Commun*, 10, 5029. https://doi.org/10.1038/s41467-019-13036-1.

Kassambara, A., 2023a. ggpubr: 'ggplot2' Based Publication Ready Plots. <https://CRAN.R-project.org/package=ggpubr>

Kassambara, A., 2023b. rstatix: Pipe-Friendly Framework for Basic Statistical Tests. <https://CRAN.R-project.org/package=rstatix>

Kaunisto, K. M., Roslin, T., Sääksjärvi, I. E., Vesterinen, E. J., 2017. Pellets of proof: First glimpse of the dietary composition of adult odonates as revealed by metabarcoding of feces. *Ecol Evol*, 7, 8588–8598. https://doi.org/10.1002/ECE3.3404.

Kaunisto, K. M., Roslin, T., Forbes, M. R., Morrill, A., Sääksjärvi, I. E., Puisto, A. I. E., Lilley, T. M., Vesterinen, E. J., 2020. Threats from the air: Damselfly predation on diverse prey taxa. *J Anim Ecol*, 89, 1365–1374. https://doi.org/10.1111/1365-2656.13184.

Kaur, M., Karnwal, A., 2023. Screening of endophytic Bacteria from stress-tolerating plants for abiotic stress tolerance and plant growth-promoting properties: Identification of potential strains for bioremediation and crop enhancement. *J Agric Food Res*, 14, 100723. https://doi.org/10.1016/J.JAFR.2023.100723.

Koboldt, D. C., Steinberg, K. M., Larson, D. E., Wilson, R. K., Mardis, E. R., 2013. The next-generation sequencing revolution and its impact on genomics. *Cell*, 155, 38. https://doi.org/10.1016/j.cell.2013.09.006.

Koga, R., Meng, X. Y., Tsuchida, T., Fukatsu, T., 2012. Cellular mechanism for selective vertical transmission of an obligate insect symbiont at the bacteriocyte-embryo interface. *Proc Natl Acad Sci U S A*, 109, E1230–E1237. https://doi.org/10.1073/pnas.1119212109.

Kogawa, M., Nishikawa, Y., Saeki, T., Yoda, T., Arikawa, K., Takeyama, H., Hosokawa, M., 2023. Revealing within-species diversity in uncultured human gut bacteria with single-cell long-read sequencing. *Front Microbiol*, 14, 1133917. https://doi.org/10.3389/fmicb.2023.1133917.

Kreer, C., Döring, M., Lehnen, N., Ercanoglu, M. S., Gieselmann, L., Luca, D., Jain, K., Schommers, P., Pfeifer, N., Klein, F., 2020. openPrimeR for multiplex amplification of highly diverse templates. *J Immunol Methods*, 480, 112752. https://doi.org/10.1016/J.JIM.2020.112752.

Kristensen, D. M., Wolf, Y. I., Koonin, E. V., 2017. ATGC database and ATGC-COGs: an updated resource for micro- and macro-evolutionary studies of prokaryotic genomes and protein family annotation. *Nucleic Acids Res*, 45, D210. https://doi.org/10.1093/NAR/GKW934.

Kuchina, A., Brettner, L. M., Paleologu, L., Roco, C. M., Rosenberg, A. B., Carignano, A., Kibler, R., Hirano, M., DePaolo, R. W., Seelig, G., 2021. Microbial single-cell RNA sequencing by split-pool barcoding. *Science*, 371. https://doi.org/10.1126/science.aba5257.

Labouyrie, M., Ballabio, C., Romero, F., Panagos, P., Jones, A., Schmid, M. W., Mikryukov, V., Dulya, O., Tedersoo, L., Bahram, M., Lugato, E., van der Heijden, M. G. A., Orgiazzi, A., 2023. Patterns in soil microbial diversity across Europe. *Nat Commun*, 14, 3311. https://doi.org/10.1038/s41467-023-37937-4.

Lan, F., Demaree, B., Ahmed, N., Abate, A. R., 2017. Single-cell genome sequencing at ultra-high-throughput with microfluidic droplet barcoding. *Nat Biotechnol*, 35, 640–646. https://doi.org/10.1038/nbt.3880.

Lange, C., Boyer, S., Bezemer, T. M., Lefort, M. C., Dhami, M. K., Biggs, E., Groenteman, R., Fowler, S. V., Paynter, Q., Verdecia Mogena, A. M., Kaltenpoth, M., 2023. Impact of intraspecific variation in insect microbiomes on host phenotype and evolution. *ISME J*, 17, 1798–1807. https://doi.org/10.1038/s41396-023-01500-2.

Langille, M. G. I., Zaneveld, J., Caporaso, J. G., McDonald, D., Knights, D., Reyes, J. A., Clemente, J. C., Burkepile, D. E., Vega Thurber, R. L., Knight, R., Beiko, R. G., Huttenhower, C., 2013. Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nat Biotechnol*, 31, 814–821. https://doi.org/10.1038/nbt.2676.

Leonaviciene, G., Leonavicius, K., Meskys, R., Mazutis, L., 2020. Multi-step processing of single cells using semi-permeable capsules. *Lab Chip*, 20, 4052–4062. https://doi.org/10.1039/D0LC00660B.

Leviatan, S., Shoer, S., Rothschild, D., Gorodetski, M., Segal, E., 2022. An expanded reference map of the human gut microbiome reveals hundreds of previously unknown species. *Nat Commun*, 13, 3863. https://doi.org/10.1038/s41467-022-31502-1.

Li, H., Shu, X., Zhou, J., Meng, L., Zhou, X., Obrycki, J. J., Li, B., 2022. A comparison of microbiota in the ladybird Harmonia axyridis between native and invaded ranges. *Entomol Exp Appl*, 170, 993–1000. https://doi.org/10.1111/eea.13230.

de Lima, J. D., Monteiro, P. H. R., Rivadavea, W. R., Barbosa, M., Cordeiro, R. D., Garboggini, F. F., Auer, C. G., da Silva, G. J., 2024. Potential of endophytic bacteria from Acacia mearnsii: Phosphate solubilization, indole acetic acid production, and application in wheat. *Appl Soil Ecol*, 196, 105315. https://doi.org/10.1016/J.APSOIL.2024.105315.

Lin, H., Peddada, S. Das., 2020. Analysis of compositions of microbiomes with bias correction. *Nat Commun*, 11, 3514. https://doi.org/10.1038/s41467-020-17041-7.

Ling, N., Wang, T., Kuzyakov, Y., 2022. Rhizosphere bacteriome structure and functions. *Nat Commun*, 13, 836. https://doi.org/10.1038/s41467-022-28448-9.

Liu, T., Zhao, H., Wang, T., 2020. An empirical Bayes approach to normalization and differential abundance testing for microbiome data. *BMC Bioinformatics*, 21, 225. https://doi.org/10.1186/s12859-020-03552-z.

Lloréns-Rico, V., Simcock, J. A., Huys, G. R. B., Raes, J., 2022. Single-cell approaches in human microbiome research. *Cell*, 185, 2725–2738. https://doi.org/10.1016/J.CELL.2022.06.040.

Lobanov, A., Dyckman, S., Kurkjian, H., Momeni, B., 2023. Spatial structure favors microbial coexistence except when slower mediator diffusion weakens interactions. *Elife*, 12, e82504. https://doi.org/10.7554/ELIFE.82504.

Losey, J. E., Vaughan, M., 2006. The Economic Value of Ecological Services Provided by Insects. *Bioscience*, 56, 311–323. https://doi.org/10.1641/0006-3568(2006)56[311:TEVOES]2.0.CO;2.

Ma, L. chao, Zhao, H. qing, Wu, L. B., Cheng, Z. le, Liu, C., 2023. Impact of the microbiome on human, animal, and environmental health from a One Health perspective. *Sci One Health*, 2, 100037. https://doi.org/10.1016/J.SOH.2023.100037.

Manogaran, M., Ahmad, S. A., Yasid, N. A., Yakasai, H. M., Shukor, M. Y., 2018. Characterisation of the simultaneous molybdenum reduction and glyphosate degradation by Burkholderia vietnamiensis AQ5-12 and Burkholderia sp. AQ5-13. *3 Biotech*, 8, 117. https://doi.org/10.1007/s13205-018-1141-2.

Martin, M., 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J*, 17, 10–12. https://doi.org/10.14806/ej.17.1.200.

Mathew, S. A., Muola, A., Saikkonen, K., Saloniemi, I., Helander, M., Puigbò, P., 2022. Quantification of the Potential Impact of Glyphosate-Based Products on Microbiomes. *J Vis Exp*. https://doi.org/10.3791/63109.

Mathew, S. A., Fuchs, B., Nissinen, R., Helander, M., Puigbò, P., Saikkonen, K., Muola, A., 2023. Glyphosate-based herbicide use affects individual microbial taxa in strawberry endosphere but not the microbial community composition. *J Appl Microbiol*, 134, lxad006. https://doi.org/10.1093/jambio/lxad006.

Mathew, S. A., Jeevannavar, A., Helander, M., Tamminen, M., Puigbò, P., Rainio, M., Saikkonen, K., Fuchs, B., 2024. Glyphosate residues in soil and phosphate fertilizer affect foliar endophytic microbial community composition and phytohormone levels in potato. *Phytobiomes J*, 8. https://doi.org/10.1094/PBIOMES-11-23-0118-R.

McNulty, R., Sritharan, D., Pahng, S. H., Meisch, J. P., Liu, S., Brennan, M. A., Saxer, G., Hormoz, S., Rosenthal, A. Z., 2023. Probe-based bacterial single-cell RNA sequencing predicts toxin regulation. *Nat Microbiol*, 8, 934–945. https://doi.org/10.1038/s41564-023-01348-4.

Miller, M. B., Bassler, B. L., 2001. Quorum sensing in bacteria. *Annu Rev Microbiol*, 55, 165–199. https://doi.org/10.1146/ANNUREV.MICRO.55.1.165.

Montagna, M., Chouaia, B., Mazza, G., Prosdocimi, E. M., Crotti, E., Mereghetti, V., Vacchini, V., Giorgi, A., De Biase, A., Longo, S., Cervo, R., Lozzia, G. C., Alma, A., Bandi, C., Daffonchio, D., 2015. Effects of the Diet on the Microbiota of the Red Palm Weevil (Coleoptera: Dryophthoridae). *PLoS One*, 10, e0117439. https://doi.org/10.1371/JOURNAL.PONE.0117439.

Moran, N. A., Russell, J. A., Koga, R., Fukatsu, T., 2005. Evolutionary Relationships of Three New Species of Enterobacteriaceae Living as Symbionts of Aphids and Other Insects. *Appl Environ Microbiol*, 71, 3302. https://doi.org/10.1128/AEM.71.6.3302-3310.2005.

Moran, N. A., McCutcheon, J. P., Nakabachi, A., 2008. Genomics and evolution of heritable bacterial symbionts. *Annu Rev Genet*, 42, 165–190. https://doi.org/10.1146/annurev.genet.41.110306.130119.

Morrill, A., Forbes, M. R., Vesterinen, E. J., Tamminen, M., Sääksjärvi, I. E., Kaunisto, K. M., 2023. Molecular Characterisation of Faecal Bacterial Assemblages Among Four Species of Syntopic Odonates. *Microb Ecol*, 87, 16. https://doi.org/10.1007/S00248-023-02328-1.

Myers, J. P., Antoniou, M. N., Blumberg, B., Carroll, L., Colborn, T., Everett, L. G., Hansen, M., Landrigan, P. J., Lanphear, B. P., Mesnage, R., Vandenberg, L. N., Vom Saal, F. S., Welshons, W. V., Benbrook, C. M., 2016. Concerns over use of glyphosate-based herbicides and risks associated with exposures: A consensus statement. *Environ Health*, 15, 19. https://doi.org/10.1186/s12940-016-0117-0.

Nielsen, A. T., Tolker-Nielsen, T., Barken, K. B., Molin, S., 2000. Role of commensal relationships on the spatial structure of a surface-attached microbial consortium. *Environ Microbiol*, 2, 59–68. https://doi.org/10.1046/J.1462-2920.2000.00084.X.

Nishikawa, Y., Kogawa, M., Hosokawa, M., Wagatsuma, R., Mineta, K., Takahashi, K., Ide, K., Yura, K., Behzad, H., Gojobori, T., Takeyama, H., 2022. Validation of the application of gel beads-based single-cell genome sequencing platform to soil and seawater. *ISME Commun*, 2, 92. https://doi.org/10.1038/S43705-022-00179-4.

Nobles, S., Jackson, C. R., 2020. Effects of Life Stage, Site, and Species on the Dragonfly Gut Microbiome. *Microorganisms*, 8, 183. https://doi.org/10.3390/MICROORGANISMS8020183.

Oksanen, J., Simpson, G. L., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., O'Hara, R. B., Solymos, P., Stevens, M. H. H., Szoecs, E., Wagner, H., Barbour, M., Bedward, M., Bolker, B., Borcard, D., Carvalho, G., Chirico, M., De Caceres, M., Durand, S., Evangelista, H. B. A., FitzJohn, R., Friendly, M., Furneaux, B., Hannigan, G., Hill, M. O., Lahti, L., McGlinn, D., Ouellette, M.-H., Ribeiro Cunha, E., Smith, T., Stier, A., Ter Braak, C. J. F., Weedon, J., 2022. vegan: Community Ecology Package. <https://CRAN.R-project.org/package=vegan>

Pan, J., Huang, T., Yao, F., Huang, Z., Powell, C. A., Qiu, S., Guan, X., 2008. Expression and characterization of aiiA gene from Bacillus subtilis BS-1. *Microbiol Res*, 163, 711–716. https://doi.org/10.1016/J.MICRES.2007.12.002.

Pande, S., Shitut, S., Freund, L., Westermann, M., Bertels, F., Colesie, C., Bischofs, I. B., Kost, C., 2015. Metabolic cross-feeding via intercellular nanotubes among bacteria. *Nat Commun*, 6, 6238. https://doi.org/10.1038/ncomms7238.

Pande, S., Kaftan, F., Lang, S., Svato, A., Germerodt, S., Kost, C., 2016. Privatization of cooperative benefits stabilizes mutualistic cross-feeding interactions in spatially structured environments. *ISME J*, 10, 1413–1423. https://doi.org/10.1038/ISMEJ.2015.212.

Petrushin, I. S., Filinova, N. V, Gutnik, D. I., 2024. Potato Microbiome: Relationship with Environmental Factors and Approaches for Microbiome Modulation. *Int J Mol Sci*, 25, 750. https://doi.org/10.3390/ijms25020750.

Pfeilmeier, S., Petti, G. C., Bortfeld-Miller, M., Daniel, B., Field, C. M., Sunagawa, S., Vorholt, J. A., 2021. The plant NADPH oxidase RBOHD is required for microbiota homeostasis in leaves. *Nat Microbiol*, 6, 852–864. https://doi.org/10.1038/s41564-021-00929-5.

Pierce, E. C., Dutton, R. J., 2022. Putting microbial interactions back into community contexts. *Curr Opin Microbiol*, 65, 56–63. https://doi.org/10.1016/j.mib.2021.10.008.

Poirier, Y., Bucher, M., 2002. Phosphate Transport and Homeostasis in Arabidopsis. *Arabidopsis Book*, 1, e0024. https://doi.org/10.1199/TAB.0024.

R Core Team., 2025. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>

Rahme, L. G., Stevens, E. J., Wolfort, S. F., Shao, J., Tompkins, R. G., Ausubel, F. M., 1995. Common Virulence Factors for Bacterial Pathogenicity in Plants and Animals. *Science*, 268, 1899–1902. https://doi.org/10.1126/SCIENCE.7604262.

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. *Environ Microbiol Rep*, 13, 309–316. https://doi.org/10.1111/1758-2229.12931.

Ramula, S., Mathew, S. A., Kalske, A., Nissinen, R., Saikkonen, K., Helander, M., 2022. Glyphosate residues alter the microbiota of a perennial weed with a minimal indirect impact on plant performance. *Plant Soil*, 472, 161–174. https://doi.org/10.1007/s11104-021-05196-1.

Ratnasingham, S., Hebert, P. D. N., 2007. bold: The Barcode of Life Data System (http://www.barcodinglife.org). *Mol Ecol Notes*, 7, 364. https://doi.org/10.1111/J.1471-8286.2007.01678.X.

Ren, Z., Cai, T., Wan, Y., Zeng, Q., Li, C., Zhang, J., Ma, K., He, S., Li, J., Wan, H., 2023. Unintended consequences: Disrupting microbial communities of Nilaparvata lugens with non-target pesticides. *Pestic Biochem Physiol*, 194, 105522. https://doi.org/10.1016/J.PESTBP.2023.105522.

Richardson, M., Zhao, S., Lin, L., Sheth, R. U., Qu, Y., Lee, J., Moody, T., Ricaurte, D., Huang, Y., Velez-Cortes, F., Urtecho, G., Wang, H. H., 2025. SAMPL-seq reveals micron-scale spatial hubs in the human gut microbiome. *Nat Microbiol*, 10, 527–540. https://doi.org/10.1038/s41564-024-01914-4.

Rico-Jiménez, M., Roca, A., Krell, T., Matilla, M. A., 2022. A bacterial chemoreceptor that mediates chemotaxis to two different plant hormones. *Environ Microbiol*, 24, 3580–3597. https://doi.org/10.1111/1462-2920.15920.

Rizzi, A., Crotti, E., Borruso, L., Jucker, C., Lupi, D., Colombo, M., Daffonchio, D., 2013. Characterization of the bacterial community associated with larvae and adults of anoplophora chinensis collected in Italy by culture and culture-independent methods. *Biomed Res Int*. https://doi.org/10.1155/2013/420287.

Roberts, D. W., 2023. labdsv: Ordination and Multivariate Analysis for Ecology. <https://CRAN.R-project.org/package=labdsv>

Rognes, T., Flouri, T., Nichols, B., Quince, C., Mahé, F., 2016. VSEARCH: A versatile open source tool for metagenomics. *PeerJ*, 4, e2584. https://doi.org/10.7717/PEERJ.2584.

Rosenberg, A. B., Roco, C. M., Muscat, R. A., Kuchina, A., Sample, P., Yao, Z., Graybuck, L. T., Peeler, D. J., Mukherjee, S., Chen, W., Pun, S. H., Sellers, D. L., Tasic, B., Seelig, G., 2018. Single-cell profiling of the developing mouse brain and spinal cord with split-pool barcoding. *Science*, 360, 176–182. https://doi.org/10.1126/science.aam8999.

Posit team., 2025. RStudio: Integrated Development Environment for R. Posit Software, PBC, Boston, MA. <http://www.posit.co/>

Ruuskanen, S., Fuchs, B., Nissinen, R., Puigbò, P., Rainio, M., Saikkonen, K., Helander, M., 2023. Ecosystem consequences of herbicides: the role of microbiome. *Trends Ecol Evol*, 38, 35–43. https://doi.org/10.1016/j.tree.2022.09.009.

Ryan, R. P., Monchy, S., Cardinale, M., Taghavi, S., Crossman, L., Avison, M. B., Berg, G., van der Lelie, D., Dow, J. M., 2009. The versatility and adaptation of bacteria from the genus Stenotrophomonas. *Nat Rev Microbiol*, 7, 514–525. https://doi.org/10.1038/nrmicro2163.

Sabree, Z. L., Kambhampati, S., Moran, N. A., 2009. Nitrogen recycling and nutritional provisioning by Blattabacterium, the cockroach endosymbiont. *Proc Natl Acad Sci U S A*, 106, 19521–19526. https://doi.org/10.1073/pnas.0907504106.

Saqib, H. S. A., Sun, L., Pozsgai, G., Liang, P., Goraya, M. U., Akutse, K. S., You, M., Gurr, G. M., You, S., 2023. Gut microbiota assemblages of generalist predators are driven by local- and landscape-scale factors. *Front Microbiol*, 14, 1172184. https://doi.org/10.3389/fmicb.2023.1172184.

Schönbrunn, E., Eschenburg, S., Shuttleworth, W. A., Schloss, J. V., Amrhein, N., Evans, J. N. S., Kabsch, W., 2001. Interaction of the herbicide glyphosate with its target enzyme 5-enolpyruvylshikimate 3-phosphate synthase in atomic detail. *Proc Natl Acad Sci U S A*, 98, 1376–1380. https://doi.org/10.1073/pnas.98.4.1376.

Schreiber, F., Littmann, S., Lavik, G., Escrig, S., Meibom, A., Kuypers, M. M. M., Ackermann, M., 2016. Phenotypic heterogeneity driven by nutrient limitation promotes growth in fluctuating environments. *Nat Microbiol*, 1, 16055. https://doi.org/10.1038/NMICROBIOL.2016.55.

Seemann, T., 2018. Barrnap, BAsic Rapid Ribosomal RNA Predictor. <https://github.com/tseemann/barrnap>

Sharon, G., Segal, D., Zilber-Rosenberg, I., Rosenberg, E., 2011. Symbiotic bacteria are responsible for diet-induced mating preference in Drosophila melanogaster, providing support for the hologenome concept of evolution. *Gut Microbes*, 2, 190–192. https://doi.org/10.4161/GMIC.2.3.16103.

Sheik, A. R., Muller, E. E. L., Audinot, J. N., Lebrun, L. A., Grysan, P., Guignard, C., Wilmes, P., 2016. In situ phenotypic heterogeneity among single cells of the filamentous bacterium Candidatus Microthrix parvicella. *ISME J*, 10, 1274–1279. https://doi.org/10.1038/ISMEJ.2015.181.

Shen, Y., Fan, N., Ma, S. xia, Cheng, X., Yang, X., Wang, G., 2025. Gut Microbiota Dysbiosis: Pathogenesis, Diseases, Prevention, and Therapy. *MedComm*, 6, e70168. https://doi.org/10.1002/MCO2.70168.

Sheth, R. U., Li, M., Jiang, W., Sims, P. A., Leong, K. W., Wang, H. H., 2019. Spatial metagenomic characterization of microbial biogeography in the gut. *Nat Biotechnol*, 37, 877–883. https://doi.org/10.1038/s41587-019-0183-2.

Shi, H., Shi, Q., Grodner, B., Lenz, J. S., Zipfel, W. R., Brito, I. L., De Vlaminck, I., 2020. Highly multiplexed spatial mapping of microbial communities. *Nature*, 588, 676–681. https://doi.org/10.1038/s41586-020-2983-4.

Shyu, S. J., Lee, R. C. T., 1990. Solving the set cover problem on a supercomputer. *Parallel Comput*, 13, 295–300. https://doi.org/10.1016/0167-8191(90)90132-S.

Singh, S., Kumar, V., Datta, S., Wani, A. B., Dhanjal, D. S., Romero, R., Singh, J., 2020. Glyphosate uptake, translocation, resistance emergence in crops, analytical monitoring, toxicity and

degradation: a review. *Environ Chem Lett*, 18, 663–702. https://doi.org/10.1007/s10311-020-00969-z.

Smith, C. C., Srygley, R. B., Healy, F., Swaminath, K., Mueller, U. G., 2017. Spatial structure of the mormon cricket gut microbiome and its predicted contribution to nutrition and immune function. *Front Microbiol*, 8, 801. https://doi.org/10.3389/fmicb.2017.00801.

Sobol, M. S., Kaster, A. K., 2023. Back to Basics: A Simplified Improvement to Multiple Displacement Amplification for Microbial Single-Cell Genomics. *Int J Mol Sci*, 24, 4270. https://doi.org/10.3390/IJMS24054270.

Spencer, S. J., Tamminen, M. V., Preheim, S. P., Guo, M. T., Briggs, A. W., Brito, I. L., A Weitz, D., Pitkänen, L. K., Vigneault, F., Virta, M. P., Alm, E. J., 2016. Massively parallel sequencing of single cells by epicPCR links functional genes with phylogenetic markers. *ISME J*, 10, 427–436. https://doi.org/10.1038/ismej.2015.124.

Spudich, J. L., Koshland, D. E., 1976. Non-genetic individuality: Chance in the single cell. *Nature*, 262, 467–471. https://doi.org/10.1038/262467a0.

Straub, D., Blackwell, N., Langarica-Fuentes, A., Peltzer, A., Nahnsen, S., Kleindienst, S., 2020. Interpretations of Environmental Microbial Community Studies Are Biased by the Selected 16S rRNA (Gene) Amplicon Sequencing Pipeline. *Front Microbiol*, 11, 550420. https://doi.org/10.3389/fmicb.2020.550420.

Stromberg, Z. R., Phillips, S. M. B., Omberg, K. M., Hess, B. M., 2023. High-throughput functional trait testing for bacterial pathogens. *mSphere*, 8, e00315-23. https://doi.org/10.1128/MSPHERE.00315-23.

Su, P., Kang, H., Peng, Q., Wicaksono, W. A., Berg, G., Liu, Z., Ma, J., Zhang, D., Cernava, T., Liu, Y., 2024. Microbiome homeostasis on rice leaves is regulated by a precursor molecule of lignin biosynthesis. *Nat Commun*, 15, 23. https://doi.org/10.1038/s41467-023-44335-3.

Suliasih, Widawati, S., 2021. Inorganic and organic phosphate solubilization potential of Stenotrophomonas maltophilia. *IOP Conf Ser Earth Environ Sci*, 948, 12054. https://doi.org/10.1088/1755-1315/948/1/012054.

Tamminen, M., Spaak, J., Caduff, L., Schiff, H., Lang, R., Schmid, S., Montealegre, M. C., Julian, T. R., 2020. Digital multiplex ligation assay for highly multiplexed screening of β-lactamase-encoding genes in bacterial isolates. *Commun Biol*, 3, 264. https://doi.org/10.1038/S42003-020-0980-7.

Tang, W., Li, Q., Chen, L., Zhang, W. X., Wang, H., 2021. Biofilm Community Structures and Opportunistic Pathogen Gene Markers in Drinking Water Mains and the Role of Pipe Materials. *ACS ES&T Water*, 1, 630–640. https://doi.org/10.1021/ACSESTWATER.0C00137.

Tedersoo, L., Albertsen, M., Anslan, S., Callahan, B., 2021. Perspectives and Benefits of High-Throughput Long-Read Sequencing in Microbial Ecology. *Appl Environ Microbiol*, 87, e00626-21. https://doi.org/10.1128/AEM.00626-21.

Theys, C., Verheyen, J., Delnat, V., Janssens, L., Tüzün, N., Stoks, R., 2023. Thermal and latitudinal patterns in pace-of-life traits are partly mediated by the gut microbiome. *Sci Total Environ*, 855, 158829. https://doi.org/10.1016/J.SCITOTENV.2022.158829.

Tokuda, G., Watanabe, H., 2007. Hidden cellulases in termites: Revision of an old hypothesis. *Biol Lett*, 3, 336–339. https://doi.org/10.1098/rsbl.2007.0073.

Veening, J. W., Stewart, E. J., Berngruber, T. W., Taddei, F., Kuipers, O. P., Hamoen, L. W., 2008. Bet-hedging and epigenetic inheritance in bacterial cell development. *Proc Natl Acad Sci U S A*, 105, 4393–4398. https://doi.org/10.1073/PNAS.0700463105.

Vesterinen, E. J., Kaunisto, K. M., Lilley, T. M., 2020. A global class reunion with multiple groups feasting on the declining insect smorgasbord. *Sci Rep*, 10, 16595. https://doi.org/10.1038/s41598-020-73609-9.

Virgo, M., Mostowy, S., Ho, B. T., 2025. Emerging models to study competitive interactions within bacterial communities. *Trends Microbiol*, 33, 688–700. https://doi.org/10.1016/J.TIM.2024.12.009.

Walter, M., Frampton, C. M., Boyd-Wilson, K. S. H., Harris-Virgin, P., Waipara, N. W., 2007. Agrichemical impact on growth and survival of non-target apple phyllosphere microorganisms. *Can J Microbiol*, 53, 45–55. https://doi.org/10.1139/W06-093.

Wang, X., Sun, S., Yang, X., Cheng, J., Wei, H., Li, Z., Michaud, J. P., Liu, X., 2020. Variability of Gut Microbiota Across the Life Cycle of Grapholita molesta (Lepidoptera: Tortricidae). *Front Microbiol*, 11, 1366. https://doi.org/10.3389/fmicb.2020.01366.

Wang, Y., Ma, L., Liu, Z., Chen, J., Song, H., Wang, J., Cui, H., Yang, Z., Xiao, S., Liu, K., An, L., Chen, S., 2022. Microbial interactions play an important role in regulating the effects of plant species on soil bacterial diversity. *Front Microbiol*, 13, 984200. https://doi.org/10.3389/fmicb.2022.984200.

Wang, B., Lin, A. E., Yuan, J., Novak, K. E., Koch, M. D., Wingreen, N. S., Adamson, B., Gitai, Z., 2023. Single-cell massively-parallel multiplexed microbial sequencing (M3-seq) identifies rare bacterial populations and profiles phage infection. *Nat Microbiol*, 8, 1846–1862. https://doi.org/10.1038/s41564-023-01462-3.

Wang, Y., Rozen, D. E., 2018. Gut microbiota in the burying beetle, Nicrophorus vespilloides, provide colonization resistance against larval bacterial pathogens. *Ecol Evol*, 8, 1646–1654. https://doi.org/10.1002/ece3.3589.

Wisniewski-Dyé, F., Vial, L., 2008. Phase and antigenic variation mediated by genome modifications. *Antonie Van Leeuwenhoek*, 94, 493–515. https://doi.org/10.1007/s10482-008-9267-6.

Wistrand-Yuen, E., Knopp, M., Hjort, K., Koskiniemi, S., Berg, O. G., Andersson, D. I., 2018. Evolution of high-level resistance during low-level antibiotic exposure. *Nat Commun*, 9, 1599. https://doi.org/10.1038/s41467-018-04059-1.

Wong, A. C. N., Wang, Q. P., Morimoto, J., Senior, A. M., Lihoreau, M., Neely, G. G., Simpson, S. J., Ponton, F., 2017. Gut Microbiota Modifies Olfactory-Guided Microbial Preferences and Foraging Decisions in Drosophila. *Curr Biol*, 27, 2397-2404.e4. https://doi.org/10.1016/j.cub.2017.07.022.

Xu, N., Qu, Q., Zhang, Z., Yuan, W., Cui, H., Shen, Y., Lin, W., Lu, T., Qian, H., 2020. Effects of residual S-metolachlor in soil on the phyllosphere microbial communities of wheat (Triticum aestivum L.). *Sci Total Environ*, 748, 141342. https://doi.org/10.1016/J.SCITOTENV.2020.141342.

Yang, L., Chen, J., 2022. A comprehensive evaluation of microbial differential abundance analysis methods: current status and potential solutions. *Microbiome*, 10, 130. https://doi.org/10.1186/s40168-022-01320-0.

Yang, L. H., Gratton, C., 2014. Insects as drivers of ecosystem processes. *Curr Opin Insect Sci*, 2, 26–32. https://doi.org/10.1016/J.COIS.2014.06.004.

Yanni, D., Márquez-Zacarías, P., Yunker, P. J., Ratcliff, W. C., 2019. Drivers of Spatial Structure in Social Microbial Communities. Curr Biol, 29, R545–R550. https://doi.org/10.1016/j.cub.2019.03.068.

Yarza, P., Yilmaz, P., Pruesse, E., Glöckner, F. O., Ludwig, W., Schleifer, K.-H., Whitman, W. B., Euzéby, J., Amann, R., Rosselló-Móra, R., 2014. Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. *Nat Rev Microbiol*, 12, 635–645. https://doi.org/10.1038/nrmicro3330.

Zangenberg, G., Saiki, R. K., Reynolds, R., 1999. Multiplex PCR: Optimization guidelines. PCR Applications: Protocols for Functional Genomics (pp. 73–94). Academic Press, Alameda, California.

Zhang, J., Liu, Y. X., Zhang, N., Hu, B., Jin, T., Xu, H., Qin, Y., Yan, P., Zhang, X., Guo, X., Hui, J., Cao, S., Wang, X., Wang, C., Wang, H., Qu, B., Fan, G., Yuan, L., Garrido-Oter, R., Chu, C., Bai, Y., 2019. NRT1.1B is associated with root microbiota composition and nitrogen use in field-grown rice. *Nat Biotechnol*, 37, 676–684. https://doi.org/10.1038/s41587-019-0104-4.

Zheng, B., Zhao, Q., Feng, L., Zhang, Z., Zhang, Q., Zhang, F., Deng, Y., Lu, T., Qian, H., 2022. Regulative effect of imazethapyr on Arabidopsis thaliana growth and rhizosphere microbial

community through multiple generations of culture. *Plant Soil*, 473, 625–637. https://doi.org/10.1007/s11104-022-05318-3.

Zhuang, Y., Liu, S., Xiao, J., Chen, T., Gao, D., Xu, Y., Jiang, W., Wang, J., Hou, G., Li, S., Zhao, X., Huang, Y., Li, S., Zhang, S., Li, M., Wang, W., Li, S., Cao, Z., 2024. Metagenomics reveals the characteristics and potential spread of microbiomes and virulence factor genes in the dairy cattle production system. *J Hazard Mater*, 480, 136005. https://doi.org/10.1016/J.JHAZMAT.2024.136005.

Zilionis, R., Nainys, J., Veres, A., Savova, V., Zemmour, D., Klein, A. M., Mazutis, L., 2016. Single-cell barcoding and sequencing using droplet microfluidics. *Nat Protoc*, 12, 44–73. https://doi.org/10.1038/nprot.2016.154.

TURUN
YLIOPISTO
UNIVERSITY
OF TURKU

Painosalama, Turku, Finland 2026

ISBN 978-952-02-0501-0 (PRINT)
ISBN 978-952-02-0502-7 (PDF)
ISSN 0082-6979 (Print)
ISSN 2343-3183 (Online)