

# **Microbiome assembly, dynamics, and recruitment within the wheat root**

**Samuel M.M. Prudence**

Doctor of philosophy

**Department of Molecular Microbiology**

**John Innes Centre**

**&**

**School of Biological Sciences**

**University of East Anglia**

This copy of the thesis has been supplied on condition that anyone who consults it is understood to recognise that its copyright rests with the author and that use of any information derived there from must be in accordance with current UK Copyright Law. In addition, any quotation or extract must include full attribution.

## Abstract

Wheat is a staple crop for 40% of the global population. However, yields over the remainder of the 21<sup>st</sup> century will become strained by climate change, necessitating new innovations to maintain and increase productivity. Root associated microbial communities have demonstrated the capacity to improve yields by increasing nutrient bioavailability, alleviating abiotic stress, and providing disease protection.

This project aimed to characterise the microbial community associated with wheat, to identify core microbial taxa associated with the roots, thus likely to provide benefits to the host. This project also aimed to understand which factors influence the microbiome, and which of these taxa utilise host derived carbon.

16S rRNA gene and ITS2 region metabarcoding of the bacterial, fungal, and archaeal communities within the rhizosphere and endosphere of wheat revealed that soil type had a major impact on the community composition, whilst plant genotype had a limited effect on the microbiome. Five core bacterial families were enriched within the rhizosphere or endosphere of wheat regardless of soil type or genotype, *Streptomyetaceae*, *Burkholderiaceae*, *Pseudomonadaceae*, *Rhizobiaceae*, and *Chitinophagaceae*. *Streptomyetaceae* and *Burkholderiaceae* were the most abundant families within the endosphere. Full length 16S rRNA gene sequencing resolved these groups to the species or genus level. Developmental senescence was shown to negatively impact the abundance of these groups, demonstrating input from the living plant is required to maintain their presence within the endosphere. Stable isotope probing showed nine bacterial taxa utilised host derived carbon, including *Pseudomonadaceae* and *Burkholderiaceae*.

Overall, this project has provided significant progress towards our understanding of the core bacterial families associated with wheat roots. This can be followed up with investigations into the roles these microbes play within the root, and how they interact with the host. In the future this understanding could lead to new ways of utilising the capabilities of the microbial community for agriculture.

## **Access Condition and Agreement**

Each deposit in UEA Digital Repository is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of the Data Collections is not permitted, except that material may be duplicated by you for your research use or for educational purposes in electronic or print form. You must obtain permission from the copyright holder, usually the author, for any other use. Exceptions only apply where a deposit may be explicitly provided under a stated licence, such as a Creative Commons licence or Open Government licence.

Electronic or print copies may not be offered, whether for sale or otherwise to anyone, unless explicitly stated under a Creative Commons or Open Government license. Unauthorised reproduction, editing or reformatting for resale purposes is explicitly prohibited (except where approved by the copyright holder themselves) and UEA reserves the right to take immediate 'take down' action on behalf of the copyright and/or rights holder if this Access condition of the UEA Digital Repository is breached. Any material in this database has been supplied on the understanding that it is copyright material and that no quotation from the material may be published without proper acknowledgement.

## Contents

<b>Abstract</b> .....	<b>2</b>
<b>Contents</b> .....	<b>3</b>
<b>Figure Index</b> .....	<b>6</b>
<b>Table Index</b> .....	<b>8</b>
<b>Acknowledgements</b> .....	<b>9</b>
<b>Publications Arising From This Work</b> .....	<b>11</b>
<b>Chapter 1. Introduction</b> .....	<b>12</b>
1.1    Challenges facing agriculture in a changing climate .....	12
1.2    Plant root microbiomes, function, dynamics, and assembly .....	15
1.2.1    Microbiome dynamics .....	16
1.2.2    Microbiome assemblies .....	18
1.2.3    Stable isotope probing for the identification of root exudate utilising microorganisms .....	33
1.2.4    Challenges and opportunities for studying the endosphere microbiome.....	42
1.2.5    The wheat seed microbiome and its influence on root microbial communities .....	43
1.2.6    Developmental senescence and its potential influence on plant root microbiomes .....	45
1.2.7    Archaea within plant root microbiomes .....	46
1.3    Project aims.....	52
<b>Chapter 2. Materials and methods</b> .....	<b>53</b>
2.1    Chemicals and reagents .....	56
2.2    Environmental sampling .....	56
2.2.1    Agricultural soil sampling and chemical analysis.....	56
2.2.2    Field wheat sampling.....	57
2.3    Plant cultivation, sampling, and experimental methods.....	58
2.3.1    Wheat seed surface sterilisation & cultivation .....	58
2.3.2    Sterile wheat seedling sampling .....	58

2.3.3	Stable Isotope labelling of wheat root exudates using $^{13}\text{CO}_2$ .....	58
2.3.4	Wheat root sampling and experimental methods .....	60
2.3.5	Archaeal plant growth promotion experiments .....	61
2.4	Nucleic acid methods.....	63
2.4.1	Electrophoresis.....	63
2.4.2	Amplification and sequencing of 16S rRNA gene and ITS2 sequences from root and soil DNA extracts .....	64
2.4.3	End-point PCR screening for <i>Burkholderiaceae</i> family endophytic bacteria	65
2.4.4	Real time quantitative PCR.....	66
2.4.5	Purification of PCR products and plasmids .....	70
2.4.6	Density gradient ultracentrifugation.....	70
2.5	Microbiological techniques.....	73
2.5.1	Endophyte spot bioactivity screen.....	73
2.5.2	Test for root endophytes interactions with <i>Streptomyces</i> .....	73
2.5.3	Wheat take-all fungus bioassays .....	73
2.5.4	Ammonia oxidising archaea enrichment culture.....	74
2.6	Computational and statistical analysis .....	74
2.6.1	Illumina amplicon sequencing analysis .....	74
2.6.2	PacBio full length 16S rRNA gene sequencing analysis .....	76

### **Chapter 3. Cross-domain compositional analysis of the wheat root**

<b>microbiome .....</b>	<b>79</b>	
3.1	Profiling the microbiome associated with field grown wheat.....	79
3.2	The effect of developmental senescence on the wheat root microbiome .	87
3.3	Archaea within the wheat root microbiome .....	90
3.3.1	The efficacy of short-read amplicon sequencing for archaeal diversity studies.....	90
3.3.2	Assessing plant-growth promoting potential of an ammonia oxidising archaeon .....	91
3.4	Discussion and conclusions.....	94

<b>Chapter 4. Drivers of microbial diversity within wheat roots .....</b>	<b>100</b>
4.1 Pot grown wheat as an agriculturally relevant model .....	100
4.2 Soil as a primary driver of root community diversity .....	104
4.3 Is <i>Triticum aestivum</i> var. Paragon an outlier among UK elite spring bread wheat? .....	109
4.4 The endophytic seed microbiome .....	118
4.5 Long read amplicon sequencing for genus to species level identification of core endosphere bacteria.....	120
4.6 Discussion and conclusions.....	131
<b>Chapter 5. Identification of root exudate utilising microbes within the wheat root microbiome using DNA stable isotope probing (SIP).....</b>	<b>143</b>
5.1 Archaeal 16S rRNA gene SIP.....	144
5.2 Fungal 18S ITS2 region SIP .....	144
5.3 Bacterial 16S rRNA gene SIP .....	147
5.3.1 CO <sub>2</sub> fixation by soil autotrophs.....	147
5.3.2 Endosphere compartment .....	147
5.3.3 Rhizosphere compartment.....	151
5.4 Discussion and conclusions.....	155
<b>Chapter 6. Isolation and characterisation of <i>Burkholderiaceae</i> and <i>Pseudomonas</i> endophytes .....</b>	<b>165</b>
6.1 Isolation and identification of root endophytes .....	165
6.2 Bioactivity of isolates .....	167
6.3 Testing strains for interaction with <i>Streptomyces</i> .....	169
6.4 Assessment of phytase activity from isolates .....	173
6.5 Discussion .....	173
<b>Chapter 7. Conclusions and future work.....</b>	<b>175</b>
7.1 Conclusions.....	175
7.2 Future work .....	177
<b>Bibliography.....</b>	<b>182</b>
<b>Supplementary Information.....</b>	<b>203</b>

## Figure index

<b>Figure 1.1.</b> Root microbiome compartments.....	17
<b>Figure 1.2.</b> Stable isotope probing from plants.....	3
<b>Figure 3.1.</b> Amplification of archaeal 16S rRNA gene sequences from wheat roots.....	80
<b>Figure 3.2.</b> PCoA for bacteria, archaeal, and fungal diversity across root compartments for stem elongation and senescent growth phases.....	83
<b>Figure 3.3.</b> Community barplots for bacteria, archaeal, and fungi across root compartments for stem elongation and senescent growth phases.....	84
<b>Figure 3.4.</b> qPCR for bacterial, archaeal, and fungal abundance at stem elongation and senescent growth phases .....	85
<b>Figure 3.5.</b> Differential abundance showing enrichment of fungal or bacterial taxa at stem elongation and senescent growth phases.....	86
<b>Figure 3.6.</b> Differential abundance comparing the endosphere abundance of fungal or bacterial taxa at stem elongation or senescent growth phases.....	88
<b>Figure 3.7.</b> DGGE for archaeal 16S rRNA gene or <i>amoA</i> gene diversity across compartments.....	92
<b>Figure 3.8.</b> <i>Nitrosocosmicus franklandus</i> plant growth promotion assay.....	93
<b>Figure 4.1.</b> PCoA for bacteria, archaeal, and fungal diversity across root compartments comparing pot and field grown plants, and soil types.....	102
<b>Figure 4.2.</b> Community barplots for bacteria, archaeal, and fungi across root compartments in different soil types.....	103
<b>Figure 4.3.</b> Differential abundance showing core taxa enriched .....	108
<b>Figure 4.4.</b> qPCR for bacterial, archaeal and fungal abundance in compost and agricultural soil .....	109
<b>Figure 4.5.</b> Bacterial community barplots for five varieties of wheat.....	112
<b>Figure 4.6.</b> Pedigree of UK wheat varieties.....	113
<b>Figure 4.7.</b> PCoA for the bacterial community for different wheat varieites.....	114
<b>Figure 4.8.</b> qPCR for microbial abundance within seeds.....	117
<b>Figure 4.9.</b> Amplification of microbial 16S rRNA gene sequences from the seed.....	117
<b>Figure 4.10.</b> Community barplots for the bacterial and fungal seed community....	118
<b>Figure 4.11.</b> Phylogenetic tree of <i>Burkholderiaceae</i> endosphere sequences.....	124
<b>Figure 4.12.</b> Phylogenetic tree of <i>Streptomycetaceae</i> endosphere sequence.....	125
<b>Figure 4.13.</b> Phylogenetic tree of <i>Pseudomonadaceae</i> endosphere sequences..	126
<b>Figure 4.14.</b> Phylogenetic tree of <i>Rhizobiaceae</i> endosphere sequences.....	127
<b>Figure 4.15.</b> Phylogenetic tree of <i>Chitinophageaceae</i> endosphere sequences....	128

<b>Figure 4.16.</b> Phylogenetic trees for exudate utiliser endosphere sequences.....	130
<b>Figure 5.1.</b> Amplification of archaeal and fungal barcodes from SIP fraction,.....	145
<b>Figure 5.2.</b> DGGE for archaeal community diversity in SIP fractions.....	145
<b>Figure 5.3.</b> qPCR for archaeal 16S rRNA gene abundance within SIP fractions..	146
<b>Figure 5.4.</b> qPCR for fungal 18S rRNA gene abundance within SIP fractions .....	146
<b>Figure 5.5.</b> DGGE for bacterial community in endosphere SIP fractions.....	148
<b>Figure 5.6.</b> Community barplots and PCoA for the bacterial community within endosphere SIP fractions.....	149
<b>Figure 5.7.</b> DGGE and qPCR for bacterial community abundance and diversity across rhizosphere SIP fractions.....	150
<b>Figure 5.8.</b> PCoA for bacterial community across rhizosphere SIP fractions.....	151
<b>Figure 5.9.</b> Abundance and differential abundance for exudate utilising bacteria.	154
<b>Figure 6.1.</b> <i>Pseudomonas</i> Take-All bioassay.....	170
<b>Figure 6.2.</b> Bioassay Screen.....	171
<b>Figure 6.2.</b> <i>Streptomyces</i> interaction assay.....	172
<b>Supplementary Figure S.1.</b> <i>Nitrosocosmicus franklandus</i> <i>Arabidopsis</i> plant growth promotion experiment.....	203
<b>Supplementary Figure S.2.</b> <i>Nitrosocosmicus franklandus</i> enrichment culture in plant growth medium.....	204
<b>Supplementary Figure S.3.</b> qPCR experiment for bacterial abundance within sterile seedling roots.....	204
<b>Supplementary Figure S.4.</b> Community barplots for the unfractionated community of plants used in the SIP experiment.....	205
<b>Supplementary Figure S.5-S.13.</b> Quality plots for PacBio Sequencing data.....	205
<b>Supplementary Figure S.14-S.15.</b> Melt curves for qPCR assays.....	215

## Table index

<b>Table 1.1.</b> Wheat root microbiome studies .....	23
<b>Table 1.2.</b> Archaea plant interaction studies.....	49
<b>Table 2.1.</b> Media recipes.....	53
<b>Table 2.2.</b> Buffer recipes.....	55
<b>Table 2.3.</b> Soil chemical properties.....	57
<b>Table 2.4.</b> Primer sequences.....	67
<b>Table 2.5.</b> PCR conditions.....	69
<b>Table 2.6.</b> SIP fractions for sequencing.....	72
<b>Table 2.7.</b> Qiime2 Dada2 settings.....	76
<b>Table 2.8.</b> PacBio sequencing analysis stats.....	78
<b>Table 3.1.</b> SIMPER outputs for field grown wheat.....	82
<b>Table 3.2.</b> Permanova results, senescent/stem elongation.....	89
<b>Table 3.3.</b> Dunns test results, AOA plant growth promotion.....	93
<b>Table 3.1.</b> SIMPER outputs for field grown wheat.....	82
<b>Table 4.1.</b> Permanova results pot/field and soil type.....	104
<b>Table 4.2.</b> SIMPER outputs for soil type.....	105
<b>Table 4.3.</b> Pairwise permanova for wheat varieties.....	115
<b>Table 4.4.</b> ANCOM-BC results comparing varieties endosphere.....	115
<b>Table 4.5.</b> ANCOM-BC results comparing varieties rhizosphere.....	115
<b>Table 4.6.</b> ANCOM-BC results comparing compartments.....	116
<b>Table 5.1.</b> Differential abundance analysis identifying CO <sub>2</sub> -fixing autotrophs.....	147
<b>Table 6.1.</b> Isolate strain list.....	167
<b>Supplementary Table S.1.</b> Qiime2 taxonomy based filtering stats.....	217
<b>Supplementary Table S.2-S.8.</b> DESeq2 outputs.....	223
<b>Supplementary Table S.9-S.15.</b> Identity matrices from PacBio sequencing.....	232

## Acknowledgements

I would firstly like to thank my supervisory team, Professor Matt Hutchings, Dr Laura Lehtovirta-Morley, and Professor Colin Murrell, for the opportunity to undertake this research project and to pursue this degree. In particular I would like to thank Matt for taking me into his research group, for introducing me to the incredible world of *Streptomyces* biology, and for offering day to day support and advice as we navigated this research project. I will be forever grateful for the advice and guidance Matt has given me over the past five years, not only for conducting scientific research, but for helping me develop academic soft skills, and for the help and advice he has given me for establishing the next phase of my career. I would also like to thank Colin for lending me his expertise within the field of environmental microbiology, and his invaluable advice on many of the experiments described in this thesis. Further, I would like to thank him for allowing me to use the facilities within his laboratory, and for the help he has given me to establish the next stages of my academic career. I would also like to give my thanks to Laura for introducing me to the amazing field of archaeal biology, and for all of the advice, technical training, and guidance she has given me for both the study of archaea, and within the environmental microbiology field generally. I would also like to thank her for allowing me to use the facilities in her laboratory, for her advice on navigating academia, and for the support and advice she has given me for the next stages of my career.

I would also like to offer my thanks to all the members of the Hutchings lab who have helped me through this project. In particular huge thanks go out to Dr Sarah Worsley for all the technical training, advice, and guidance she gave me throughout this project. Without both the scientific groundwork she laid during her time in the Hutchings lab, and the training and guidance she gave me throughout this project, this work would not have been possible.

Huge thanks go out to Dr Jake Newitt, with whom I worked very closely on the stable isotope probing experiments, and whose training and advice was invaluable. As Jake and I collaborated closely on the DNA stable isotope probing experiment we split the substantial labour of the manual pulse  $^{13}\text{CO}_2$  injections, and of the density gradient ultracentrifugation, 50:50. I then assumed responsibility for the fungal and archaeal community profiling whilst Jake performed DGGE, qPCR, DNA sequencing, and data analysis for the bacterial community, and this data has also been presented within his thesis. The data presented by figures 5.6, 5.7, and 5.8

was produced by Jake, in addition to tables 2.6 and 5.1. Jake also gave me the 16S rRNA gene sequences for the *Streptomyces* isolates used in the long-read sequencing analysis presented in chapter 4, section 4.5. Jake also lent me the *Streptomyces* CRS3 endosphere isolate used for experiments presented in chapter six. I would like to extend my gratitude to Jake for being an open and collaborative scientific colleague throughout this project, for sharing the load of performing SIP, and for the invaluable scientific discussions we have had over the years.

I would also like to offer my thanks to Dr Neil Holmes, who trained me with basic microbiology methods when I first started as an undergraduate, and then with *Streptomyces* genetics when I began my PhD. I would like to extend my thanks to George Pawley, who assisted me with the isolation and characterisation of endosphere isolates during his undergraduate research project, and who isolated and sequenced some of the strains presented in chapter six. I would also like to extend my thanks to the other members of the Hutchings lab who have offered me support, training, and companionship throughout these years, Dr Rebecca Devine, Dr Tom McLean, Dr Nicolle Som, and all the rest.

I would like to thank Dr Gregory Rix for performing phytase degradation assays on the isolates described in supplementary chapter seven. Further I would like to thank Greg and all the members of the Pamela Salter Office at University of East Anglia (UEA), the Molecular Microbiology department at the John Innes Centre (JIC), and my cohort of the EnvEast doctoral training program for their companionship and support.

I would also like to thank Simon Orford and the whole team at the JIC Germplasm Resources Unit, and the JIC field studies team. Throughout this project the GRU supplied the wheat seeds needed to perform these experiments, and both Simon and the field studies site team were instrumental in granting us access to the field site, and to helping us with the sampling of plants and soil that underpinned much of this project. Further I express thanks to Professor Cristobal Uauy for his advice on the experimental design of experiments investigating the influence of genotype on microbiome composition. I would also like to thank Elaine Patrick of the research support staff at the UEA for her technical support during my time there. I would like to thank the UEA high performance computing team for their support enabled the computational analysis presented in this thesis. I also extend my thanks to the staff of the EnvEast doctoral training program for all their work to support me and the rest of the cohort.

Lastly, I am grateful to all of my friends and family for their support throughout these years. Particularly to all of my friends in the Sorrow discord server, and to my parents and my sisters.

## **Publications arising from this work**

Prudence, S. M. M.\* , Newitt, J. T.\* , Worsley, S. F., Macey, M. C., Murrell, J. C., Lehtovirta-Morley L. E. & Hutchings, M. I. (2021) Soil, senescence, and exudate utilisation: characterisation of the Paragon var. spring bread wheat root microbiome. *Environmental Microbiome* **16**, 12. DOI: 10.1186/s40793-021-00381-2

Newitt, J. T.\* , Prudence, S. M. M.\* , Hutchings, M. I. & Worsley, S. F. (2019). Biocontrol of Cereal Crop Diseases Using Streptomycetes. *Pathogens* **8**, 78. DOI: 10.3390/pathogens8020078

Prudence, S. M. M., Addington, E.\* , Castaño, L.\* , Mark, D. M.\* , Pintor-Escobar, L.\* , Russel, A. H.\* & McLean, T. C. (2020). Advances in actinomycete research: an ActinoBase review of 2019. *Microbiology* **166**, 8. DOI: 10.1099/mic.0.000944

Worsley, S. F., Macey, M. C., Prudence, S. M. M., Wilkinson, B., Murrell, J. C. & Hutchings, M. I. (2021). Investigating the role of root exudates in recruiting *Streptomyces* bacteria to the *Arabidopsis thaliana* root microbiome. *Frontiers in Molecular Biosciences* **8**, DOI: 10.1101/2020.09.09.290742.

\* These authors contributed equally

# Chapter 1. Introduction

## 1.1 Challenges facing agriculture in a changing climate

Humanity began cultivating crops between 11,000 and 23,000 years ago <sup>1,2</sup>. The earliest evidence for agriculture indicated people may have been cultivating perennial grasses, such as early progenitors of modern barley, wheat, and oat. This evidence indicates farming began 23,000 years ago within the fertile crescent, a strip of land east of the Mediterranean encompassing modern Israel and Palestine, Lebanon, and Syria <sup>2</sup>. It is more generally accepted that agricultural practices had been developed by around 11,000 years ago <sup>1</sup>. Since then, agricultural practices have been refined and developed; 21<sup>st</sup> century mechanised farming contrasts starkly to the early subsistent or feudal systems. Most pertinently agriculture was revolutionised by innovations during the green revolution of the 20<sup>th</sup> century such as the Haber Bosch process and resulting synthetic nitrogen fertilisers, selective breeding, pesticides, and mechanisation. The green revolution has precipitated immense growth in the human population and improvements in both living standards and nutrition worldwide <sup>3</sup>. While overwhelmingly beneficial, the intensification of agriculture does not however come without costs, and has presented a number of fundamental issues such as soil degradation <sup>4-6</sup>, greenhouse gas emissions <sup>7</sup>, and biodiversity loss <sup>8</sup>, which must be addressed during the 21<sup>st</sup> century as agriculture shifts to sustainable practices.

As the climate changes, and as humanity continues to alter the ecology of the biosphere, several new challenges face agriculture; these can be broadly grouped into three categories, extreme weather, soil degradation, and pestilence. On the weather front, less predictable seasonal weather, more severe droughts, and regular flooding all mean that crops must be more resilient to avoid regular catastrophic yield losses. These effects however are already being felt; over the past 5-10 years farmers in the UK have reported broad impacts on costs and yields resulting from extreme weather including heavy rainfall, droughts, extremes of temperature and more frequent storms and floods <sup>9</sup>. Heavy rainfall can have direct impacts including an increased disease incidence within the grain, and soil erosion and waterlogged fields that cause yield losses and, in some cases, complete crop loss. More frequent stormy weather can cause additional direct crop damage, and both heavy rainfall and stormy weather can cause logistical and operational problems which reduce the efficiency of running a farm. Drier weather and extreme heat can cause further operational difficulties, for example with drilling fields as dry

harder soil is more difficult to seed. This can result in poorly established crops with lower yields, or crops that are completely lost. Crops might also have greater levels of pestilence and disease, and sometimes the quality is so low that harvested produce must be rejected. Drought is particularly bad for yields of wheat (*Triticum aestivum*), which is a staple crop for more than 4 billion people and globally accounts for more than 20% of human calorie and protein consumption<sup>10</sup>. Whilst the sector has been slow to adapt, there are also few practical solutions to these issues, and many farmers are relying on business efficiencies, infrastructure improvements, and income diversification to remain financially viable<sup>9</sup>. Globally malnourishment remains high, in 2015 this was estimated to be ~10% of the global population, or 780 million people<sup>11</sup>. Extreme weather will continue to effect agriculture in the UK, and have similar effects around the world, if not more extreme in areas vulnerable to tropical storms or coastal storm surges. This underlies the need to select different species or generate new varieties of crops, or to develop new management strategies, that are resilient to extreme unseasonal conditions like drought, coastal flooding, or heavy rain.

Underlying these climactic issues are the problems rooted within the soil. As more land is used for agriculture more of the soil is eroded; on agricultural land soil production rates (which are driven by organic matter deposition) are outstripped by soil erosion caused by rainfall on the bare fallow, over time this decreases agricultural productivity<sup>4</sup>. Further, the fertility of agricultural soils is degrading as a result of intense farming<sup>5,6</sup>, necessitating the use of inorganic fertilisers to boost yields. This is a wasteful process and is contributing to the dangerous dysregulation of the global nitrogen cycle; 83% of inorganic nitrogen used for agriculture is lost to the environment, causing severe ecological problems such as eutrophication and hypoxic zones within waterways. Loss of this inorganic N to the atmosphere as potent greenhouse gases like N<sub>2</sub>O or other reactive NO<sub>x</sub> species also contributes ~6% of global greenhouse gas emissions<sup>12</sup>.

Similarly, agricultural soils are also deficient in bioavailable phosphate. Naturally, bioavailable phosphate within soil would be maintained through decaying plant matter. In agricultural fields however much of this is removed as the plant is harvested for consumption, thus nutrient levels in the soil have degraded over the centuries<sup>13</sup>. Since the 19<sup>th</sup> century, rock phosphate has been mined for use in inorganic fertilisers to increase soil fertility, this is not a renewable source of phosphorus however, and the practice causes significant environmental issues including eutrophication<sup>14</sup>. These issues underline the need for more sustainable

and ecologically responsible means for maintaining soil fertility and preventing soil erosion, to maintain high crop yields to feed a growing global population.

Pesticides are another key element of modern agriculture; it is becoming increasingly obvious however that this practice is unsustainable and ecologically irresponsible. Pesticides can have devastating impacts on local biodiversity, particularly for key ecosystem service providers like pollinators<sup>15,16</sup>. Many of these insects are prey items for birds and mammals, so declines in these primary consumers can have an effect across multiple trophic levels. This is exacerbated further by the fact that many common pesticides bioaccumulate, these compounds are often toxic to bird and mammal species, more so at the higher doses which result from these pesticides becoming concentrated up through trophic levels<sup>17</sup>. The effects of pesticides stretch further still, the pollinator services which are disrupted by pesticides are key to the reproduction of the majority of flowering plants, and therefore to the functioning of many ecosystems, and are key to the reproductive cycle of 75% of crops<sup>16</sup>. Inevitably, as with all antibiotics, the use of pesticides also causes resistance such as the fungicide resistance that has been observed for many phytopathogenic fungi<sup>18,19</sup> or for anthelmintic compounds including nematicides like avermectin<sup>20</sup>. Many phytopathogens are also now transmitted globally, and cause epidemics affecting crops across continents<sup>21</sup>, for example with wheat yellow rust (*Puccinia striiformis*)<sup>22,23</sup>. This demonstrates the need for new approaches to pathogen control, to combat globally transmitted, pesticide resistant crop diseases and to find ecologically responsible means for pest management.

These issues are compounded in part by the large quantities of food wasted in western countries, and a growing global population, which is predicted to reach 9 billion globally by 2050<sup>24</sup>. This growth is being driven by economic and healthcare improvements within major African nations<sup>24</sup> and a predicted 2.4-fold increase in per capita income<sup>3</sup>, overwhelmingly demonstrating positive progress for humanity. Increasing affluence is predicted to result in a greater proportion of the global population shifting to more energy intensive diets, for example by consuming more grain-fed meat<sup>3</sup>, and shifting to more wasteful western dietary practices. As a result, it has been estimated that an increase in food production of 25-70% will be necessary to feed the world by 2050. This, compounded by constraints on agriculture from climate change and related issues of soil degradation and pestilence, demonstrates a unique challenge for the future of agriculture, and warrants a new agricultural revolution for the 21<sup>st</sup> century.

## 1.2 Plant root microbiomes, function, dynamics, and assembly

The community of microorganisms residing within root associated niches, termed in this thesis the root associate microbiome (RAM), can provide numerous benefits to the host plant. For example, RAMs can provide plants with protection from abiotic stressors such as drought<sup>25,26</sup>, flooding<sup>27</sup>, or osmotic stress from high salt concentrations within the soil<sup>28-30</sup>. These capabilities may prove invaluable for crop management as extreme weather such as drought, storm surges, and coastal flooding become more frequent and less predictable. Root microbiomes can also improve plant nutrition. *Rhizobia*, residing within the root nodules of legumes such as pea (*Pisum sativum*), fix atmospheric nitrogen and generate ammonia, providing nitrogen to the host plant<sup>31</sup>. Microbes within the soil are also responsible for the solubilisation and subsequent bioavailability of other important nutrients such as phosphate<sup>32-34</sup>, or important minerals or metals such as iron<sup>35,36</sup>. One study has observed that co-inoculation of a symbiotic fungus and bacterium had a positive effect on soy bean (*Glycine max*) root uptake of magnesium, manganese, iron, potassium, calcium, copper, zinc, boron, and sulphur<sup>36</sup>. One of the most well-known examples are arbuscular mycorrhizal fungi (AMF) which can promote host health via phosphate solubilisation<sup>34</sup>. This demonstrates the vital role that RAMs could play in the remediation of soils, and in sustainable crop fertilisation strategies.

Many root-associated microbiota have also demonstrated the capacity to impede the growth of fungal or bacterial pathogens. Bacteria such as *Pseudomonas* spp. and *Streptomyces* spp. are responsible for the generation of disease suppressive soils<sup>37-40</sup>, where the survival of pathogenic fungi is inhibited, thus protecting plants cultivated within those fields from disease. Many groups have also been shown to protect from disease directly *in planta*, example genera include *Streptomyces*<sup>41,42</sup>, *Paenibacillus*<sup>43</sup>, or *Pseudomonas*<sup>40,44</sup>. A number of microbial biocontrol formulations are already available commercially, for example Mycostop® for which the active agent is an inoculum of *Streptomyces griseoviridis* K61<sup>45</sup>, or Actinovate® which contains *Streptomyces lydicus* WYEC 108<sup>46</sup>. Both these formulations are advertised as antifungal biocontrol products, and both strains have a demonstrated capacity to promote plant growth in greenhouse experiments<sup>47,48</sup>. Microbial biocontrol, utilising the capacity of the RAM, can also be explored as an alternative to harmful chemical pesticides. All together the members of root microbiome can have significant benefits for host plants and represent a potential source of new microbial biotechnologies for crop management, and to address the issues such as climate change which are facing agriculture in the 21<sup>st</sup> century.

### 1.2.1 Microbiome dynamics

Microbial communities can be stochastic, and while the evidence indicates that plant roots select for specific microbial lineages<sup>49-52</sup>, RAMs are no exception to this stochasticity. This presents a unique challenge for microbial biotechnologies aiming to apply the capabilities of the microbiome. A variety of factors influence the composition of soil and root-associated microbial communities, and thus are likely to affect the efficacy of proposed microbial interventions designed for agriculture. These factors lie within two broad categories, those which are driven by the host plant and those which are determined by the environment. The most important environmental factor is soil type<sup>52-56</sup>, and a number of factors contribute to defining soil type. The soil pH exerts the most important influence upon the microbial community<sup>53</sup>, along with the level of essential nutrients such as inorganic nitrogen, soluble phosphate, potassium, and magnesium. Other important soil characteristics include the organic matter content, the water content, and the quantities of trace metals (both contaminating metals such as cadmium<sup>57</sup> and biologically useful metals like iron<sup>35,58</sup>). The most commonly used soil categories refer to the soil's structure and mechanical properties, as many papers describe soil using terms such as sandy loam, clay, silt, or compost<sup>53,56,59-61</sup>. This refers to the soils structural and mechanical properties. Soil structure will influence the oxygenation of the soil, which in turn massively impacts microbial activity within the soil, and root microbiome composition<sup>56,60</sup>.

Whilst soil is the most important environmental factor influencing the function and composition of microbial communities within plant roots, other environmental factors will influence plant associated microbial communities. These include the weather and climate (and climate change)<sup>11,62</sup>, and farming practices such as irrigation, fertilisation, pesticide use, tillage, and pre-cropping/crop rotations<sup>63-67</sup>. Plant-driven factors also have a huge influence on microbial communities, these include the plant species<sup>50,56,59,68</sup>, and genetic variation within species<sup>50,51,69-72</sup>. Further, plant metabolism and development can have a massive impact through variable root exude profiles<sup>69,73,74</sup>, which plants alter throughout life to select for and maintain specific bacterial lineages within the RAM. All of these interacting factors mean that reliably modelling the microbiome for a given crop across different sites is extremely challenging, and further complicated by the fact that many of these factors may vary significantly each growing season<sup>62,75,76</sup>.

The root associated microbial community is highly complex, and consists of a diverse community of bacteria, archaea, fungi, and other micro-eukarya such as nematodes and protists <sup>77</sup>. Given prevalence of fungi and bacteria within this community, it stands to reason that bacteriophage and mycovirus, both prevalent within soils <sup>78,79</sup>, would also be a significant component of the root community. One study identified positive selection for phage defence genes within the barley (*Hordeum vulgare*) rhizosphere <sup>51</sup>, supporting the idea that phage impart a significant selection pressure upon the rhizosphere bacterial community. Interestingly bacteriophages can interact with fungal mycelia <sup>80</sup>, it is thought that this could increase phage retention within soils and contribute to phage transport via fungal mycelia. Whilst little is known about mycoviruses within root communities, they have been identified infecting both fungal pathogens <sup>81,82</sup> and fungal symbionts <sup>83</sup>, demonstrating that they are indeed a biological force acting within the root microbiome. These examples underline the extraordinary complexity of the root associated microbial community, and the broad diversity of biological functions and niches that must be accounted for to fully comprehend the ecology of this system.

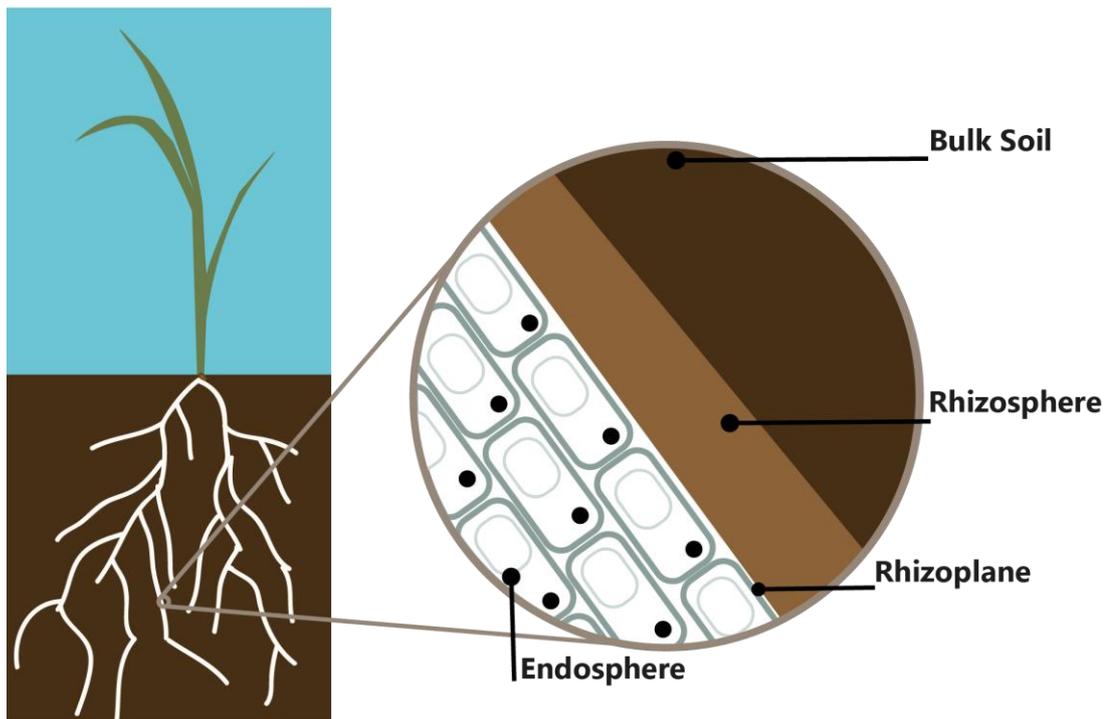


Figure 1.1 Diagram demonstrating the four major compartments of the root associated microbiome, the bulk soil, rhizosphere, rhizoplane and endosphere compartments. Black dots within the plant cells represent the nucleus, and the rounded grey shape represents the vacuole.

When studying the root microbiome, the community can be segmented into four compartments (Figure 1.1). The bulk soil describes the microbial community within the soil which is not associated with the roots. The majority of microbiota within the

RAM are recruited horizontally from the soil community <sup>52</sup>, and soil properties exert a major influence on community composition, as discussed earlier in this section. The bulk soil community is therefore an important determinant of microbiome composition within the root <sup>52,55,84</sup>, and must be considered when studying the root associated community. The rhizosphere is a thin layer of soil associated with the surface of the root, this is chemically distinct from the bulk soil due to the deposition of host-derived metabolites exuded from the root, known as root exudates <sup>85</sup>. This effect that plants have on the microbial community is known as the rhizosphere effect, and the strength and range of this effect can vary between plant species <sup>56,68</sup>. The rhizoplane describes the root surface, which is home to many bacteria and to ectomycorrhizal fungi <sup>86–88</sup>. Lastly the endosphere describes the root interior, where the most closely associated beneficial microbiota reside such as arbuscular mycorrhizal fungi (AMF), also known as endomycorrhiza <sup>70</sup>.

Within these compartments there is much complexity and variability amongst the microbial community. Resultantly, a detailed understanding of the bulk of the molecular mechanisms underpinning microbiome assembly function remains elusive. We know which microbial taxa can be present within RAMs, and we have identified many of the factors that can influence community composition. For some specific groups we also know how mechanistically they might benefit the host, for example through the production of siderophores <sup>35,89</sup> or antibiotics <sup>32,42,90</sup>, or through phosphate solubilisation <sup>32–34</sup>. How these specific mechanisms interact within the entire ecosystem, and how within the environment they combine to provide benefits for the host, remains to be fully elucidated. We still do not fully understand how observed community level changes practically influence the ecology and functionality of this microbial ecosystem due to the complexity and variability of RAMs, the network of interactions within root communities is extremely complex and therefore difficult to fully resolve.

## **1.2.2 Microbiome assemblies**

### **1.2.2.1 The core microbiome**

While RAMs contain diverse microbial taxa, spanning all domains of life, only a small subsection of this community will directly impact plant health, with the majority of the community acting as passive members of the root ecosystem. To cut through the complexity and variability of microbial communities, and to focus on those microorganisms which are most likely to directly impact the host plant, it can be

useful to determine the core microbiome. Discussed at length by Vandenkoornhuise and colleagues<sup>91</sup>, the core microbiome is defined as the microbial taxa consistently associated with a particular plant species or variety, regardless of habitat or conditions, and which provide a service either to the host plant or to the broader ecosystem<sup>92,93</sup>. This is different to the holobiont, which is a term used for studies where evolution is the primary focus; the holobiont describes an organism and all of its symbiotic microbiota as a single unit upon which selection acts<sup>91</sup>. The core microbiome however describes the subsection of the microbial niches within the RAM that provides an essential service to the host plant, or the ecosystem upon which the host plant depends. The purpose of this concept is to identify microorganisms with the greatest impact on plant health, such that the interactions of these organisms with the plant can be studied in detail. To identify these core associated taxa, approaches such as metabarcoding or metagenomics can be used to profile community diversity. To identify the core microbiome amongst this microbial diversity a number of factors must be accounted for. Soil type is one of the most important factors<sup>55,56,94</sup>, along with developmental stage<sup>95,96</sup>, genotype<sup>72,95,97</sup> and, in the case of crop plants, agricultural management strategy<sup>63,96,98,99</sup>. Once accounted for, the taxa found enriched within root associated compartments across all these conditions can be considered those most likely to constitute the core microbiome. Follow-up studies can then seek to characterise the specific interactions between these microorganisms and the host plant, and between those microorganisms and other root microbiome community members. The core microbiome of the model plant *Arabidopsis thaliana* is well studied<sup>52,100</sup>, and it has also been characterized for numerous other plant species to varying degrees<sup>101–103</sup>.

There is not a clear consensus within the literature as to how the core microbiome should be defined, and numerous approaches have been used depending on the system of study<sup>104</sup>. One limitation of these studies is that microbial community surveys are also often limited to investigations of bacterial diversity, or in some cases fungal diversity, meaning that knowledge of wheat root community diversity is limited to these two groups. Root-associated archaea, and micro-eukarya such as nematodes and protists, are considerably understudied within the RAM, archaea particularly so within terrestrial plant species. Exclusion of these groups from core microbiome studies may result in key microbial interactions or functions being overlooked.

Further to this, functional redundancy can be difficult to account for within taxonomy-focussed core microbiota studies<sup>91</sup>, as different microbial taxa are often

capable of occupying the same niche or providing the same services to the host plant. Whilst the same function may be carried out within the community, in a different environmental context a different microorganism, or multiple microorganisms, may be occupying the same niche and providing the same service. Thus, the basic Venn Diagram approach to defining the core microbiome would not identify this group as a core microbiome member despite the fact it may be providing key services to the plant. This means that studies could miss an important community function. Many approaches to defining the core microbiome also tend to disregard the rare taxa within the microbial community. It can often be challenging to distinguish these taxa from sequencing artefacts<sup>105,106</sup> meaning actual biological importance is often difficult to unravel. However, rare taxa can play important roles within the RAM<sup>106,107</sup>, and so the inability to properly assess this is a challenge for the field.

Despite these limitations, the concept of a core microbiome is a useful tool to identify microbes consistently associated with plants, and which we can hypothesize may play an important plant-beneficial role within the root ecosystem. The identification of such taxa provides focus to further studies characterising the mechanisms of specific interactions between core root microbiota and the host.

### **1.2.2.2 The core wheat root microbiome**

The microbiome associated with wheat has been investigated across a wide variety of studies<sup>63,68,94–99,108–129</sup>. Most studies focused on the rhizosphere, and have characterised the impact of a range of factors on the microbiome composition including agricultural management practice<sup>63,67,96,98,99,130</sup>, development<sup>95</sup>, land use history<sup>109</sup>, soil type<sup>115–117</sup>, and genotype<sup>72,97,108,110,111,114,128,129</sup>. The methodology used varies across these studies. Some older studies used denaturing gradient gel electrophoresis (DGGE, discussed in section 1.2.3.1)<sup>95</sup>, terminal restriction fragment length polymorphism (T-RFLP) fingerprinting<sup>112</sup>, fatty acid methyl esters (FAME), or community level physiological profiling (CLPP) to profile the rhizosphere community<sup>110,116</sup>. FAME is a gas chromatography (GC) based method for profiling microbial communities based on cell membrane fatty acid profiles, and CLPP is a culture dependant BIOLOG™ based method which profiles communities based on sole-carbon substrate utilisation. While FAME and CLPP are both able to distinguish differences in microbial community composition, they do not yield high-taxonomic resolutions and are far less accessible than the DNA, RNA, and protein-based methods common in microbial ecology.

A number of studies have used metatranscriptomics to profile the rhizosphere community<sup>68,115</sup>, and some recent studies have performed metagenomics on the rhizosphere community<sup>118,119</sup>. The vast majority of studies however used metabarcoding, sequencing barcoding regions such as the 16S rRNA gene or ITS region using Illumina platforms<sup>98,99,111,113</sup>, PacBio SMRT closed circular sequencing (CSS)<sup>127</sup>, or Roche 454 pyrosequencing<sup>94,121,123</sup>. For an overview of methods used for these studies see Table 1.1. Whilst methodology improves with time, some methodologies provide distinct benefits over others. Pyrosequencing and PacBio SMRT CSS can provide full length 16S rRNA gene or ITS region sequences; pyrosequencing however has fallen out of fashion due to the high rate of insertion and deletion errors generated when sequencing homopolymeric stretches of nucleotides, and the limited throughput when compared to Illumina sequencing<sup>131</sup>, as such this sequencing platform is no longer commonly available. With the PacBio platform, sequencing at depth can be prohibitively expensive, meaning this method is unable to provide sufficient depth at an affordable cost for many projects. Illumina sequencing uses shorter reads, thus a shorter proportion of the metabarcoding region can be sequenced so taxonomic inferences are limited to the family level<sup>132</sup>. Illumina sequencing can however affordably provide greater depth and thus better coverage for the community compared to longer read methods<sup>131,133</sup>, and is more useful for studying community dynamics. All three of these PCR based methods can introduce primer bias, and as such there is some debate regarding the accuracy of the results<sup>134,135</sup>. Differences in sequencing amplicon can also provide further opportunity for methodological difference between studies, as the region that is sequenced will affect the proportion of the total operational taxonomic units (OTUs) or amplicon sequence variants (ASV's) present that can be identified by the study<sup>132</sup>. Meta-omics approaches can alleviate a number of these problems, by sequencing all of the DNA or RNA within the rhizosphere a supposedly unbiased community profile can be produced, though factors such as DNA extraction methods can still influence results. The analysis of these datasets is also more complicated and time consuming, and performing meta-omics for the endosphere compartment is highly challenging due to contamination from host DNA sequences<sup>133</sup>. PCR based methods on the other hand can overcome this issue by amplifying a targeted metabarcoding region for the microbial organisms of interest. Whilst comparisons can be challenging, all these methods have a number of benefits and caveats, and so the most appropriate method can be selected depending on the specific biological questions of a study.

Beyond differences in sequencing approaches, a number of other methodological differences can make it difficult to compare studies. For example a range of approaches exist for defining and sampling the rhizosphere, some studies use mechanical brushing<sup>96,116</sup> whilst others wash the rhizosphere off of the roots<sup>51,137</sup>. Plant cultivation practices also vary, studies can be based in the field or based within a greenhouse or laboratory, and sampling timepoints vary by developmental stage (table 1.1). Some studies opt to pre-germinate wheat under sterile conditions<sup>68,97</sup>, whereas a number of other studies sow the plants directly into the soil. These small discrepancies between studies confound differences in the microbiome observed across studies and can make comparing the challenging.

Many studies seeking to profile the wheat microbiome have aimed to assess the impact of various different factors on community composition. The role of farming practices has been widely studied (table 1.1). Agricultural management strategy, comparing organic farming to conventional methods for example, has been shown by one study to play a relatively weak role in determining root microbiome composition<sup>99</sup>. Conversely however a more recent study found organically managed fields harbour a fungal rhizosphere community that is twice as complex as that found within conventionally managed or un-tilled fields<sup>127</sup>. This study used a longer-read sequencing methods and thus is likely to yield deeper insights; it could be that the community shifts are at the species level and so the shorter read method of the previous study was unable detect a strong fungal microbiome shift. This demonstrates how methodological differences can lead to contradictory conclusions from different studies. Cropping regimes have also been shown to have a significant impact on fungal community composition<sup>98</sup>. Other management factors that have been investigated include long term irrigation, which influenced the abundance of a broad variety of bacteria<sup>63</sup>, or long-term nitrogen fertilisation which significantly influenced microbiome composition and also reduced colonisation by *Streptomyces*<sup>96</sup>, a genus known for possessing many species with plant beneficial capabilities. Overall, these studies have demonstrated that agricultural practices can indeed have an impact on the wheat root microbiome composition and indicates that these effects should be investigated further to better inform farming practices, particularly where microbial agents are being used.

Table 1.1. Variables within a selection of papers investigating the wheat root microbiome, selected to represent the different methodologies and aims associated with wheat root microbiome research

Citation	Aims	Microbiome profiling approach	Root compartment(s)	Wheat cultivar(s)	Soil type(s)	Sampling timepoint(s)
Germida <i>et al.</i> (2001) <sup>110</sup>	Ascertain the differences between the microbiome associated with three wheat varieties and identify microorganisms responsible for these differences.	FAME and CLPP	Rhizosphere and endosphere (sampled as one)	<b>Bread:</b> PI 167549, Red fife, CDC Teal	Two similar clay loam, mildly alkaline, agricultural soils	41 days
Marschner <i>et al.</i> (2005) <sup>116</sup>	To investigate the impact of soil type and pH on rhizosphere community composition and on the role of plant and microbial mechanisms in P uptake.	FAME	Rhizosphere	<b>Bread:</b> Goldmark	A broad variety of soils sampled from 10 different sites in South Australia	42 days
Houlden <i>et al.</i> (2008) <sup>95</sup>	Assess the influence of developmental stage on bacterial and fungal rhizosphere community composition.	16S rRNA gene DGGE and CLPP	Rhizosphere	<b>Bread:</b> Pena wawa	Heavy clay, mildly alkaline, agricultural soil	12, 48, 76, 103, and 132 days
Sanguin <i>et al.</i> (2009) <sup>122</sup>	To identify shifts in the composition of the bacterial rhizosphere community during take-all decline	16S rRNA gene microarrays	Rhizosphere	<b>Bread:</b> Caphorm	Luvisoil agricultural soil	Flowering
Sachdev <i>et al.</i> (2010) <sup>125</sup>	To study the diversity of <i>Acinetobacter</i> within the wheat rhizosphere.	DGGE using an <i>Acinetobacter</i> 16S rRNA gene amplicon	Rhizosphere	<b>Bread:</b> Lokwan, HI1535, GW322	Three similar agricultural sites within the Pune district of India	45, 60, and 95 days
Hamots <i>et al.</i> (2012) <sup>124</sup>	To assess the extent of the wheat rhizosphere effect on denitrifier abundance and diversity under waterlogging stress.	<i>nirS</i> , <i>nirK</i> , and <i>nosZ</i> gene DGGE	Rhizosphere	<b>Bread:</b> Monad	Silt loam soil	92 and 105 days
Turner <i>et al.</i> (2013) <sup>68</sup>	To compare the rhizosphere microbiome community between three crop species using metatranscriptomics.	Metatranscriptomics, Pyrosesequencing	Rhizosphere	<b>Bread:</b> Paragon	One mildly alkaline agricultural soil	28 days
Yin <i>et al.</i> (2013) <sup>122</sup>	Identify microorganisms associated with <i>Rhizoctonia</i> disease suppression in the rhizosphere.	16S rRNA gene metabarcoding, Pyrosequencing	Rhizosphere	Unspecified spring bread	Silt loam agricultural soil	Unspecified

Ofek <i>et al.</i> (2014) <sup>120</sup>	To investigate the influence of plant species on bacterial rhizoplane community composition, the extent of host influence over the community composition, and to compare active (RNA) and inactive (DNA) community profiles.	16S rRNA gene and RT-16S rRNA gene metabarcoding, Pyrosequencing	Rhizoplane	<b>Durum:</b> Negev	Sandy loam agricultural soil, supplemented with Hoagland nutrient solution	12 days
Donn <i>et al.</i> (2015) <sup>112</sup>	To study whole community diversity and the dynamics of specific bacteria across multiple seasons.	16S rRNA gene T-RFLP and metabarcoding, Pyrosequencing	Rhizosphere	<b>Bread:</b> Janz, H45	Two similar loam or clay loam agricultural soils	40, 96, 126, 159, and 259 days
Tkacz <i>et al.</i> (2015) <sup>126</sup>	To investigate microbial succession within the rhizosphere microbiome.	16S rRNA gene and fungal ITS region, Pyrosequencing	Rhizosphere	<b>Bread:</b> Paragon	Two soils Compost, acidic, Agricultural soil, neutral pH	28 days
Rascovan <i>et al.</i> (2016) <sup>94</sup>	To analyse the root microbiome composition for wheat and soybean across a broad geographical area.	16S rRNA gene metabarcoding, Pyrosequencing	Rhizosphere	<b>Bread:</b> Cadenza	A broad variety of soils sampled from 11 different sites across Argentine Pampas	Unspecified
Fan <i>et al.</i> (2017) <sup>113</sup>	To assess bacterial community structure within the rhizosphere compared to the bulk soil across a broad geographical area.	16S rRNA gene metabarcoding, Illumina paired end sequencing	Rhizosphere	Unspecified bread	A broad variety of soils from nine study sites distributed over a ~800,000 km <sup>2</sup> area of the North China Plain	Unspecified
Gdanetz and Trail (2017) <sup>99</sup>	To study the stem, leaf, and root associated bacterial and fungal microbiome for wheat grown under different land management strategies.	16S rRNA gene and ITS2 region metabarcoding, Illumina paired end sequencing	Rhizosphere and endosphere (sampled as one), stem, and leaf	<b>Bread:</b> 25R39	Four similar fine loam agricultural soils	Vegetative, flowering, and seed development
Granzow <i>et al.</i> (2017) <sup>98</sup>	To investigate the influence of cropping regimes on bacterial and fungal community composition, the extent of this effect in different plant species, and if cropping regimes influence negative plant-microbe interactions.	16S rRNA gene and ITS2 region metabarcoding, Illumina paired end sequencing	Rhizosphere, endosphere, and phyllosphere	<b>Bread:</b> Hybery	Compost-sand mixture	28 days

Mahoney <i>et al.</i> (2017) <sup>111</sup>	To identify a core group of bacteria associated with the wheat rhizosphere of different wheat cultivars.	16S rRNA gene metabarcoding, Illumina paired end sequencing	Rhizosphere	<b>Bread:</b> Eltan, Finch, Hill81, Lewjain, Madsen, PI561722, PI561725, PI561726, PI561727	Two agricultural silt loam soils, mildly acidic	Approx. 9 months (overwintered)
Yin <i>et al.</i> (2017) <sup>121</sup>	To compare the rhizosphere community across multiple sites and assess the influence of tillage regimes on community composition.	16S rRNA gene metabarcoding, Pyrosequencing	Rhizosphere	Unspecified bread	Two silt loam agricultural soils, acidic	Approx. 7-8 months (overwintered)
Durán <i>et al.</i> (2018) <sup>117</sup>	To assess microbial community composition within the rhizosphere of wheat cultivated in take-all suppressive soil.	16S rRNA gene DGGE	Rhizosphere	<b>Bread:</b> Otto	Three volcanic take-all suppressive soils, pH neutral	40 days
Hayden <i>et al.</i> (2018) <sup>115</sup>	To compare the functional capacity and diversity of the rhizosphere community for wheat cultivated in <i>Rhizoctonia</i> suppressive soils to non-suppressive soils.	Metatranscriptomics, Illumina paired end sequencing	Rhizosphere	<b>Bread:</b> Gladius	Two agricultural soils from the same site in Avon, South Australia	56 days
Mavrodi <i>et al.</i> (2018) <sup>63</sup>	To investigate the influence of contrasting irrigation practices on the rhizosphere microbiome.	16S rRNA gene metabarcoding, Illumina paired end sequencing	Rhizosphere	<b>Bread:</b> Louise	Agricultural silt loam, acidic	Seven time points across the growing season
Banerjee <i>et al.</i> (2019) <sup>127</sup>	To explore the impact of farming systems on fungal community structure using PacBio SMRT sequencing.	Whole ITS region sequencing, PacBio SMRT CSS	Endosphere	<b>Bread:</b> 25 unspecified varieties	A range of different soils managed under different farming systems	Approx. 8-9 months

Chen <i>et al.</i> (2019) <sup>96</sup>	To investigate the influence of plant developmental stage and nitrogen fertilisation on the microbiome.	16S rRNA gene and fungal 18S rRNA gene metabarcoding, Illumina paired end sequencing	Rhizosphere	Undefined	Four agricultural soils with different nitrogen fertilisation regimes, mildly alkaline soils	Three time points across the growing season
Kavamura <i>et al.</i> (2019) <sup>136</sup>	To investigate the influence of land use history and microbial seed load on the root microbiome.	16S rRNA gene metabarcoding, Illumina paired end sequencing	Rhizosphere, Endosphere	<b>Bread:</b> Hereward	Bare fallow and arable soil at the Rothamstead agricultural research facility	Approx. 10 months, (overwintered)
Kuźniar <i>et al.</i> (2019) <sup>108</sup>	To identify the core endophytic microbiome.	16S rRNA gene metabarcoding, sample pooled Illumina paired end sequencing	Rhizosphere, Endosphere	<b>Bread:</b> Honda <b>Spelt:</b> Rokosz	Agricultural soil, mildly acidic	Approx. 3-4 weeks
Özkurt <i>et al.</i> (2020) <sup>129</sup>	To investigate the impact that the domestication of wheat has had on the microbiome, and the influence of seed microbiota on community composition.	16S rRNA gene and ITS1 region metabarcoding, Illumina paired end sequencing	Seed endosphere, Root endosphere, Phyllosphere	Three wild and three unspecified domesticated	Agricultural soil sampled from Germany and natural soil from Turkey, within the fertile crescent	14 days
Schlatter <i>et al.</i> (2020) <sup>109</sup>	To investigate the influence of land use history influences the root microbiome, and to identify hub taxa within the community.	16S rRNA gene and fungal ITS1 region metabarcoding, Illumina paired end sequencing	Rhizosphere	<b>Bread:</b> Louise	Four soils from farms within different rainfall zones	42 days
Simonin <i>et al.</i> (2020) <sup>128</sup>	To investigate the influence of plant genotype and soil on rhizosphere community composition in African and European soils.	16S rRNA gene and 18S rRNA gene metabarcoding, Illumina paired end sequencing	Rhizosphere	<b>Bread:</b> Apache, Bermude, Carstens, Champlain, Cheyenne, Rubisko, Soissons, Terminillo	A diverse range of eight different soils sampled from Africa and Europe, mostly silt clay or silt loam, ranging from acidic to alkaline pH	Approx. one month

<p>Tkacz <i>et al.</i> (2020) <sup>97</sup></p>	<p>To investigate the influence that selective breeding has had on the composition of the fungal and bacterial community.</p>	<p>16S rRNA gene, and both fungal and oomycete ITS1 region metabarcoding, Illumina paired end sequencing</p>	<p>Rhizosphere, Endosphere</p>	<p>Four bread, three durum, and 15 wild</p>	<p>Agricultural soil, neutral pH</p>	<p>42 days</p>
<p>Zhou <i>et al.</i> (2020) <sup>119</sup></p>	<p>To investigate the influence of decaying roots within the soil on rhizosphere microbiomes.</p>	<p>Metagenomics and 16S rRNA gene metabarcoding, paired end Illumina sequencing</p>	<p>Rhizosphere</p>	<p><b>Bread:</b> Justica CL plus</p>	<p>Agricultural soil</p>	<p>12 days</p>
<p>Iannucci <i>et al.</i> (2021) <sup>114</sup></p>	<p>To investigate the influence of root morphology, exudate composition, and genotype on microbiome composition.</p>	<p>Archaeal and bacterial 16S rRNA gene and fungal ITS1 region T-RFLP</p>	<p>Rhizosphere</p>	<p><b>Durum:</b> Cappelli, Creso, Ofanto, Simeto, Claudio, Grecale, Pedroso, PR22D89</p>	<p>60:40 soil sand mixture</p>	<p>Approx. 8-10 weeks</p>
<p>Wang <i>et al.</i> (2021) <sup>118</sup></p>	<p>To identify unculturable Zn mobilising bacteria within the rhizosphere microbiome.</p>	<p>Metagenomics, Paired end Illumina sequencing</p>	<p>Rhizosphere</p>	<p>Bei9, Xinong3517, Zhoumai24, Yannong0428, Hengguan35, Jinan17</p>	<p>Agricultural silt clay loam soil, alkaline</p>	<p>Approx. 9 months (overwintered)</p>

As discussed in sections 1.2.1 and 1.2.2.3, plant development significantly influences the root microbial community <sup>73,95</sup>, and this observation is also true for wheat <sup>95,96</sup>. Five studies have explored the influence of wheat genotype on the microbiome composition, and all of these studies used different varieties of wheat (table 1.1). For both the rhizosphere and endosphere Tkacz and colleagues found that at least half of the microbial community was conserved between a diverse range of bread and durum wheat lines, however they concluded that over generations the selective breeding of wheat had significantly altered the microbiome composition compared to wild varieties <sup>126</sup>. Iannucci and colleagues, who focused on durum wheat, also found a significant effect of genotype on the rhizosphere microbiome, corroborating the work by Tkacz and colleagues <sup>114</sup>. Similarly, Mahoney and colleagues, who focussed on the rhizosphere of several bread wheat varieties, reported a significant effect of genotype on community composition. They were also able to identify a group of core rhizosphere microbes associated with bread wheat <sup>111</sup>. Simonin and colleagues on the other hand, utilising a broad variety of soils and seven wheat genotypes, only reported a very weak effect of genotype on rhizosphere microbiome composition <sup>128</sup>. Similarly Özkurt and colleagues also found that in seedlings genotype had little effect on the root community, with very little difference between wild and domesticated varieties <sup>129</sup>. These results show that there is no clear consensus on the impact of wheat genotype on the wheat root microbiome, but in many cases a core group of microorganisms will always colonise the root compartments.

Whilst almost all of these studies used similar short read sequencing methods (table 1.1), there is one key methodological difference between these studies. For a number of studies the where microbiome composition was found to vary significantly with host genotype, that variation was driven primarily by rare or low abundance taxa; whilst some studies suggest these taxa can play important roles within the community <sup>106,107</sup>, recent discourse suggests that such taxa are likely to be sequencing artefacts <sup>105,106</sup>. Further, due to their low abundance, the measured abundance of rare taxa is vulnerable to stochastic change and thus may not be a reliable measure for analysis of compositional change. Regardless of these caveats, there is incongruence amongst the literature on this topic, and thus demand for further investigations into the influence of genotype on microbiome composition.

As discussed in section 1.2.1, soil type is amongst the most well documented factors which can influence microbial community composition. The influence of the soil on the root microbiome has been explored on a number of occasions for wheat

(table 1.1), and pH is often determined the most influential factor <sup>113,116</sup>. Despite this however Schlatter and colleagues <sup>109</sup> were still able to identify a core group of 32 taxa which colonised the wheat rhizosphere within all conditions tested, and this included families such as *Burkholderiaceae*, *Chitinophageaceae*, and *Rhizobiales*. The enrichment of these groups within the wheat rhizosphere across four different soil types strongly implies a relationship between these taxa and the plant, however the authors were only able to use a single variety of wheat within the scope of this study (*T. aestivum* var. Louise). Mahoney and colleagues <sup>111</sup> managed to identify a large core group of rhizosphere taxa within nine varieties of bread wheat. Similarly, Tkacz and colleagues <sup>97</sup> were able to identify 99 and 77 core bacterial rhizosphere or endosphere taxa respectively, associated with a diverse range of 22 different bread, durum, and wild wheat varieties. Both of these studies however chose to report these core taxa at the phyla level, making comparisons between the two studies difficult to make. Simonin and colleagues identified a core group of 85 core rhizosphere taxa, this included *Burkholderiaceae*, *Chitinophageaceae*, *Caulobacteraceae*, the fungal taxa Ascomycota (an extremely diverse group), and the cercozoan microeukaryote Filosa <sup>128</sup>. A number of these core taxa were also identified by other studies, for example Schlatter *et al.*, corroborating this finding within multiple wheat varieties and soil types, and strongly implying a general relationship between these taxa and wheat. Kuźniar *et al.* <sup>108</sup> chose to focus on the endosphere, and identified a number of core bacterial endosphere taxa including *Pseudomonas* and *Flavobacterium*. This study however focussed on a single soil type and assayed only two varieties (table 1.1). Whilst these studies represent good progress toward understanding the core rhizosphere and endosphere taxa within the wheat root microbiome, each paper has its limitations and thus no single study so far has comprehensively catalogued the core microbiome of wheat. To fully define the core microbiome for a genotypically variable and globally cultivated crop such as wheat it will take a cumulative effort from a number of studies and the whole wheat root microbiome research community, conducted on different varieties and using different soils and cultivation practices. In particular the endosphere remains understudied.

### **1.2.2.3 Root exudation and microbiome formation**

Plants exude between 20 and 40% of the carbon they fix photosynthetically as root exudate compounds <sup>138</sup> in part to modulate the microbial community associated with the roots. In this way plants can recruit beneficial microorganisms from the environment, and support their growth within the rhizosphere though

photosynthetically fixed carbon<sup>51,52,55,59,85</sup>. Root exudates are a complex mixture of organic compounds, consisting primarily of sugars, organic acids, amino acids, and fatty acids<sup>69,74,139,140</sup>. While carbon is typically the primary resource these exudates can provide to the microbial community, many of these compounds also contain nitrogen, for example amino acids and many organic acids, and some microorganisms can be supported by host derived nitrogen in addition to or instead of carbon<sup>141,142</sup>. Root exudation is a dynamic process, and varies significantly across plant species<sup>143,144</sup>, by plant growth stage<sup>73</sup>, by genotype<sup>72,114</sup>, or with root morphology<sup>143,145</sup>. Some root exudate compounds require active export, for example organic acids like citric acid and oxalic acid which require an anion pump for cell export. Genes encoding for this export machinery are primarily expressed in the root tip<sup>145</sup>. Other compounds, for example some sugars, are exported via passive diffusion across the cell membrane and thus the exudation rate correlates with root surface area. As such, the root morphology influences the exudation rates of different compounds.

Plants alter exudate composition as they develop; in early life more generally utilised resources like sugars are secreted, these non-specific substrates attract a broad range of microbiota from the soil. As the plant matures more specific compounds are secreted to maintain the presence of plant-beneficial microorganisms within root community<sup>73</sup>. Root exudation can also be altered by the colonisation of AMF<sup>146</sup>, under stress conditions<sup>147,148</sup>, by soil parameters<sup>144,149–151</sup>, or in response to agricultural soil management strategies<sup>152</sup>. For several root exudate compounds the rate of exudation is diurnal, as some compounds are primarily secreted at night during the dark, whilst others are primarily secreted during the day while the plant is photosynthesizing, or during the transition phase from dark to light<sup>150,151,153</sup>. These examples demonstrate the stochasticity of root exudation, and the complex network of inputs which regulate the composition and exudation rate of host derived metabolites into the rhizosphere.

It remains unclear how root exudates select for beneficial microbiota over neutral or pathogenic microbes. One hypothesis is that competitive exclusion is the mechanism by which pathogens are prevented from colonising the root; beneficial or neutral microorganisms colonise the root and take up niche space and resources. In doing so, they outcompete pathogenic microorganisms within the soil for these resources, as pathogens will be more specialised for plant infection over rhizosphere colonisation. In doing so, they physically exclude these organisms from the root associated community by occupying this habitat and therefore suppress their

ability to cause disease <sup>154</sup>. This hypothesis supports the general assumption that greater diversity is indicative of a healthy root microbial community. There is also however some evidence to suggest that certain molecules within the root exudates can attract specific bacterial taxa to the root <sup>51,52,85</sup>. Many bacteria exhibit chemotaxis towards plant root exudates, including taxa which include plant beneficial microorganisms such as *Pseudomonas*, *Bacillus*, *Burkholderia*, *Rhizobium*, or *Sinorhizobium* <sup>155–158</sup>. This chemotaxis can be enhanced by the presence of other non-motile members of the root community, such as AMF <sup>159</sup> or *Streptomyces* <sup>160</sup>, indicating that these microorganisms may help to recruit beneficial microbiota from the soil. Root exudates can support the growth of many microbes, for example ammonia oxidising bacterium *Nitrosolobus multiformis* isolated from the barley rhizosphere community <sup>86</sup>, a range of *Streptomyces* strains isolated from the *Arabidopsis* rhizosphere <sup>161</sup>, or AMF isolated from the roots of *Lotus japonicus*, a wild legume <sup>162</sup>.

In addition to supporting growth directly, and acting as signalling molecules for chemotaxis, root exudates can alter gene expression within root microbiota. For example, thymol within the root exudates of *Sedum alfredii* (a *Crassulaceae* family herb from China) can modulate the expression of multiple genes important for quorum sensing by *Pseudomonas aeruginosa*, acting as a quorum sensing inhibitor. Indeed, root exudates have been shown to modulate expression of a broad range of genes within *Pseudomonas*, many related to the uptake and catabolism of common root exudate compounds <sup>163</sup>. Similarly, broad changes in gene expression have been observed for rhizosphere *Bacillus* <sup>164</sup>. In response to root exudates from *Daucus carota* (wild carrot) AMF *Glomus intraradices* have been shown to alter expression of a number of genes involved germination and carbon metabolism <sup>165</sup>. These interactions can go both ways, with microorganisms altering root exudation patterns by modulating host gene expression. For example, within the peanut (*Arachis hypogaea*) rhizosphere, the fungal endophyte *Phomopsis liquidambaris* can modulate host gene expression to modulate root exudation profiles, which in turn promoted chemotaxis, growth, and biofilm formation by *Rhizobium*. This in turn increased *nodC* mediated root nodulation by this bacterium <sup>166</sup>. It is thought that the reason that *Ph. liquidambaris* modulates host gene expression in this way to promote growth and health of the host plant and the roots, thus maximising the size of its own niche.

Root exudates can also instigate changes in gene expression within pathogens, for example some plant-parasitic nematodes upregulate expression of root-infection

related genes in response to root exudates <sup>167</sup>. In the environment these root exudates do not exist in isolation however, and many microbial metabolites also interact with these effects. Mycorrhiza for example are known to also exude metabolites from roots <sup>168,169</sup>. One example demonstrated how exudates from the roots of Maize (*Zea mays*) which has been colonised by multiple AMF strains (*Funneliformis mosseae* and *Rhizophagus intraradices*) reduced expression of a mycotoxin from fungal pathogen *Fusarium proliferatum* <sup>169</sup>. The presence of AMF exudate compounds was required for the root exudates to have this effect, indicating a multipartite interaction network. Root exudates can have broad impacts on the physiology of the microbiota residing with the roots, both beneficial microbes and also on pathogenic microbes. This indicates a complicated network of interactions that can function synergistically to reduce the infectivity of pathogens or to increase the health and growth of the host plant.

While plants use root exudates to select for beneficial microbiota, they may also use them to antagonise pathogens. Some root exudate compounds exhibit antimicrobial properties, primarily effecting pathogenic microorganisms. Wheat root exudates for example contain phenolic compounds with antibacterial properties. In one study however these compounds only impacted the growth of certain bacteria, indicating that this may be a mechanism by which plants can specifically select for certain microorganisms <sup>170</sup>. In response to infection by *Fusarium*, barley begins to exude phenolic compounds with antifungal activity to repel the infection <sup>171</sup>. Among many other plant species, banana (*Musa acuminata* AAA 'Dwarf Cavendish') roots also exude antimicrobial phenolic compounds <sup>172</sup>, though common phenolic compounds are not the only antimicrobial class of root exudate. Cowpea (*Vigna unguiculata*) root exudates contain proteins which have exhibited antimicrobial properties through enzymatic activity. These chitinases are secreted to degrade the cell wall of fungal pathogens such as *Fusarium oxysporum* <sup>173</sup>. Carboxylic acids secreted from barley roots effect different members of the community to different degrees, for example having a greater inhibitory effect on *Fusarium culmorum* when compared to plant beneficial rhizobacterium *Pseudomonas fluorescens* <sup>174</sup>. As these examples demonstrate, plant root exudates can use a variety of mechanisms to selectively inhibit the growth of pathogenic microorganisms.

In summary, alongside other host elements such as immunity, root exudates are a crucial component of root microbiome assembly and as a result, research is increasingly occupied with attempting to understand the dynamics of root exudates,

to characterise their chemical composition, and to identify microorganisms who can respond to or utilise these compounds.

### **1.2.3 Stable isotope probing for the identification of root exudate utilising microorganisms**

Stable isotope probing (SIP) is a powerful tool that can be used to identify microorganisms which metabolise a particular substrate of interest, or to track the flow of isotopically labelled metabolites between microbes within a community, or between host organisms and their associated microbial communities. Thus, stable isotopes can be used in a variety of ways to provide biological insights. For example, this method has been applied extensively to identify microorganisms which degrade polluting polycyclic aromatic hydrocarbons in soils and sediments <sup>175</sup>, leading to the proposal of bioremediation strategies for contaminated sites. Recently a study used isotopically labelled amino acids to demonstrate that an asgardarchaeon, *Lokiarchaeota* MK-D1, can utilise amino acids as a growth substrate <sup>176</sup>, a novel observation for a group of organisms that had never previously been cultivated for laboratory study. In host associated systems stable isotopes have been used extensively. Using an intravenous infusion of isotopically labelled threonine and glucose, one study was able to show that host-derived carbon and nitrogen are utilised by microorganisms residing within the mouse gut <sup>177</sup>. Similarly SIP has been used to identify microorganisms associated with marine sponges which can utilise dissolved organic matter as a growth substrate <sup>178</sup>, and to characterise microbial utilisation of CO<sub>2</sub> and bicarbonate produced by rumination within the kangaroo foregut <sup>179</sup>. This shows the power that this tool has for the identification of interactions between a host and its microbial community within a diverse range of study systems.

Stable isotope probing can also be used to demonstrate the flow of nutrients more broadly within the ecosystem. In seagrass meadows chemoautotrophs residing within bivalves fix nitrogen, which is then utilised by the bivalve host for growth; SIP has demonstrated that under carbon limited conditions these chemosymbiotic bivalves then secrete excess nitrogen as ammonia, providing fertilisation for the seagrasses <sup>180</sup>. SIP can also be used to observe microbial processes within biogeochemical cycles or effecting the flow of climate active gasses. For example, SIP has been used to identify methylotrophy by microorganisms, where microbes are able to degrade the important climate active gas methane <sup>181</sup>. On the converse

SIP has also been used to identify methanogenic archaea (termed methanogens), which produce methane within rice paddy soils <sup>182</sup>. These observations could have important ramifications for strategies aiming to manage the flux of these important climate active gasses. SIP is therefore a technique with broad reaching relevance and utility and is an invaluable tool for observing biological interactions between a host and its microbiome, across trophic levels, and within important biogeochemical cycles.

Several different approaches can be used for SIP experiments depending on the aims of the investigation. For studies aiming to gain insights into microbial diversity for example DNA or RNA SIP can be used in combination with 16S rRNA gene or ITS2 sequencing. When metagenomics or metatranscriptomics are applied however DNA or RNA SIP can also be used to gain greater taxonomic resolution, and functional information about genes or transcripts used by microbes which metabolise isotopically labelled metabolites <sup>183</sup>. DNA SIP metagenomics for example been applied to the phyllosphere of oil palm (*Elaeis guineensis*), the authors were able to identify microorganisms degrading isotopically labelled isoprene, and also obtained metagenome assembled genomes (MAGs) for these bacteria, providing further insights into the potential biology of these organisms <sup>184</sup>. This in turn allowed Carrión and colleagues to identify the gene cluster within these genomes which encoded for enzymes involved in isoprene degradation, providing direct evidence of the functional capacity for isoprene degradation within isotopically labelled MAGs. Similarly, by using metatranscriptomics in conjunction with RNA SIP Dumont and colleagues were able to profile gene expression by methylotrophs active in the degradation of methane within a lake sediment <sup>185</sup>, this enabled the authors to identify the most highly expressed methane metabolism genes within the environment. These examples demonstrate the detailed information that can be obtained when SIP is applied in conjunction with other microbial ecology tools such as metagenomics or metatranscriptomics.

Protein SIP can also be used to gain functional information from SIP experiments by identifying proteins which are isotopically labelled in a metaproteomic approach <sup>186</sup>. This method is more sensitive than nucleic acid-based methods as mass spectrometry is used to detect isotope incorporation and to identify proteins. While protein SIP provides a profile of the proteins being produced by organisms metabolising isotopically labelled substrates, and thus the metabolic functions being performed by those cells, this approach does not provide detailed taxonomic identification and is often paired with DNA or RNA SIP <sup>186</sup>.

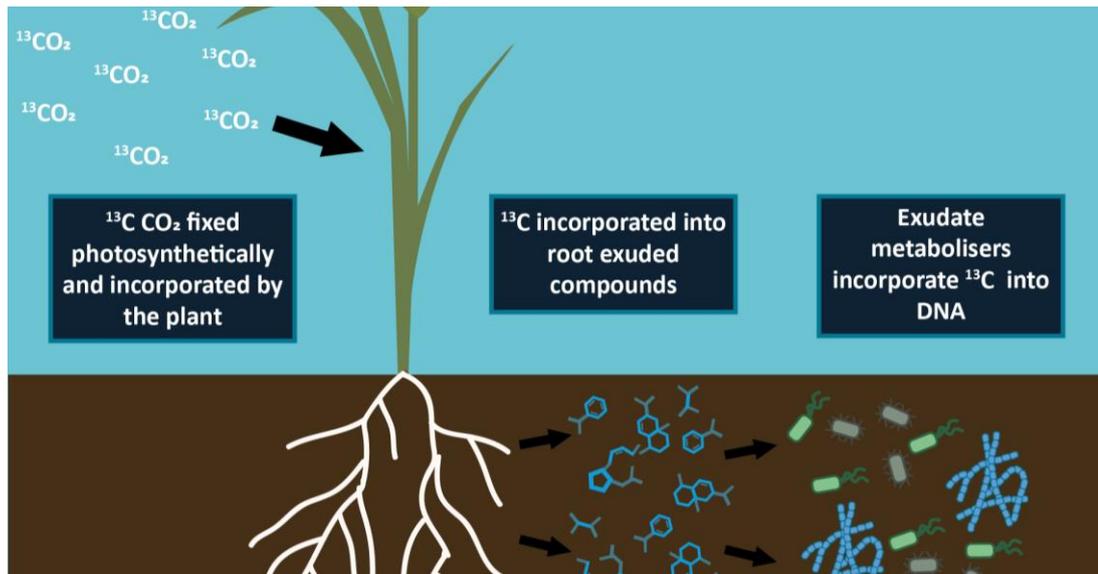


Figure 1.2 Diagram demonstrating the principle behind stable isotope probing for the identification of exudate utilising microorganisms. As the plant photosynthesizes the plant incorporates the 13-carbon from isotopically labelled  $^{13}\text{CO}_2$  into its metabolism. As such the root exudates become labelled with  $^{13}\text{C}$ , and any actively growing microorganisms within the rhizosphere which utilise these exudates as a growth substrate will incorporate  $^{13}\text{C}$  into their DNA backbone.

SIP is widely used to probe plant-microbe interactions, for example the method has been used to identify microorganisms that catabolise dead plant matter within soils<sup>187,188</sup>. Most importantly however SIP can be used to identify interactions between microbiota and living plants, for example via the identification of root-exudate utilising microbes residing within the rhizosphere (Figure 1.2)<sup>138</sup>. The most common approaches use manual  $^{13}\text{CO}_2$  pulse labelling or automatic  $^{13}\text{CO}_2$  injection. Plants are incubated in a gas tight chamber with  $\text{CO}_2$  containing a heavy stable isotope of carbon,  $^{13}\text{C}$ . As the plant photosynthesizes it fixes  $^{13}\text{CO}_2$ , and the heavy  $^{13}\text{C}$  is incorporated into the plant's metabolism. The root exudate compounds, which can constitute up to 40% of all the plants photosynthetically fixed carbon, then become labelled with  $^{13}\text{C}$  and actively growing microorganisms within the root associated community that utilise this carbon will incorporate  $^{13}\text{C}$  into their DNA backbone as they undergo genome duplication as cells divide. Total DNA can then be extracted, and density gradient ultracentrifugation and fractionation can be used to separate the heavy  $^{13}\text{C}$  labelled DNA from the  $^{12}\text{C}$  light unlabelled DNA over a caesium chloride gradient<sup>189</sup>. Root-exudate utilising organisms can then be identified via a range of different methods.

Older studies used denaturing gradient gel electrophoresis (DGGE)<sup>190</sup> or T-RFLP fingerprinting<sup>191</sup> to profile the labelled community. Most recent studies however use high throughput sequencing approaches, most commonly metabarcoding via

sequencing of the 16S rRNA gene or fungal metabarcoding genes <sup>138,183</sup>, but similar metabarcoding approaches using functional gene amplicons can be used <sup>184</sup>, as can metagenomics <sup>184</sup>. Thus, DNA SIP can be used to identify microbiota within the root community which are capable of utilising host-derived carbon, and further studies can probe the mechanisms underpinning how those microbes are interacting with the host, which exudate compounds are involved, and what effect they have on the host plant.

Numerous examples of this approach exist within the literature, and for a range of plant species cultivated under a range of conditions. Within *Arabidopsis* cultivated in compost, *Pseudomonas*, *Telluria*, *Shinella*, *Herbaspirillum*, *Sphingopyxis*, *Massilia*, and *Sinorhizobium* were all shown to utilise host derived carbon within the rhizosphere <sup>161</sup>. Another study used SIP to compare the community of exudate utilising fungi associated with four different grass species, and identified different communities of exudate-utilising fungi associated with grasses displaying different nutrient use strategies <sup>59</sup>. Several SIP studies have focussed on root exudate utilisation within the RAM of cereal crops. Rice (*Oryza sativa*) for example is widely studied; one study focussed on the identification of methanogenic archaea within the rice rhizosphere <sup>191</sup>, whilst another used SIP in conjunction with qPCR to monitor seasonal changes in methanotroph diversity <sup>192</sup>. Where other studies focussed on the archaeal community, they were able to identify a number of bacterial taxa within the rice root and the rhizosphere which were able to utilise host derived carbon <sup>193</sup>. Maize is another cereal which has been studied using SIP and DGGE <sup>194</sup>. By using SIP to probe interactions between the host plant, the bacterial community, and the hyphosphere within the maize RAM, Wang and colleagues were able to show that host derived carbon is utilised by the mycorrhizal fungus *Rhizophagus irregularis*, and subsequently transferred to a phosphate solubilising bacterium (PSB) associated with the roots, *Pseudomonas alcaligenes* <sup>195</sup>. As exemplified by these studies, when applied to plant root microbiomes SIP is a powerful tool for the identification of root-exudate utilising microorganisms, and to track the flow of host derived carbon within the microbial community.

#### **1.2.3.1 Stable isotope probing to identify exudate utilisation within wheat root microbiome**

As wheat is an important staple crop, a number of SIP studies have probed microbial metabolism within wheat associated microbiomes <sup>65,66,187,188,194,196,197</sup>. Bernard and colleagues used DNA and RNA SIP to investigate the diversity of

microorganisms utilising carbon from dead wheat tissues within the soil <sup>187</sup>. Similarly, Kaplan and colleagues aimed to assess the impact of contaminating trace metals on the ability of a soil community to degrade plant tissue, as carbon turnover is an important capability for the microbial community within healthy soils. The authors purchased <sup>13</sup>C labelled wheat roots and performed a microcosm experiment to assess the efficiency of plant tissue degradation within heavy metal contaminated soil compared to remediated soil <sup>188</sup>. This showed that after heavy metal contamination, a soil remediation project was able to restore the capacity of the soil microbial community to degrade complex plant tissues such as wheat roots. In a similar microcosm experiment Macey and colleagues aimed to identify mycorrhizal fungi within the rhizosphere of wheat <sup>197</sup>. To achieve this the authors isolated rhizosphere material from wheat and used this as the inoculum for a microcosm experiment where <sup>13</sup>C labelled methane was introduced to enable the identification of mycorrhizal fungi within the rhizosphere. Whilst these microcosm studies did not aim to investigate the root-exudate utilising community associated with wheat, they demonstrate different applications for stable isotope probing within the context of plant associated microbial science, to investigate role that soil microorganisms play in plant tissue degradation, or to study a specific subgroup of microorganisms residing within the root associated community.

Prior to the publication of the work in chapter five <sup>198</sup>, four published studies had focussed on identifying exudate utilisation within the wheat root microbiome by using stable isotope probing *in planta*. The first of these studies, from Haichar and colleagues published in 2008, used DGGE to identify exudate utilising organisms within the rhizosphere and endosphere <sup>194</sup>. To achieve this, after fractionation, 16S rRNA gene sequences were amplified from each fraction and DGGE was used. DGGE separates a heterogeneous community of DNA sequences based on the rate at which the DNA molecule denatures across a gradient of denaturant, this will depend on the GC content of the DNA sequence. Each band on a DGGE gel therefore corresponds to one unique 16S rRNA gene sequence. Haichar and colleagues identified bands that were significantly more intense, and thus likely enriched, within the heavy fractions compared to the light fractions. These bands were then excised from the gels, the DNA was purified, and the sequence was acquired via sanger sequencing. In this way Haichar and colleagues were able to identify a number of bacterial taxa enriched within the heavy fractions and hypothesized to have incorporated isotopically labelled host-derived carbon as a result of root exudate utilisation. These taxa included the *Enterobacteriaceae*,

*Xanthomonadaceae*, *Caulobacteraceae*, *Rhizobium*, *Sphingomonadaceae*, *Comamonadaceae*, *Oxalobacteraceae*, and *Paenibacillaceae*. While analysis of the statistical significance of the enrichment of these taxa was limited to computational analysis of DGGE band intensity, this study presented the first direct evidence for exudate utilisation by bacteria residing within the wheat RAM, and many subsequent studies identified similar taxa utilising root exudates within the wheat root community.

The remaining three published studies assessed root exudate utilisation within the wheat RAM using high throughput sequencing methods, which are now commonplace within contemporary microbial community ecology. These studies were from Ai *et. al.* in 2015, Uksa *et. al.* in 2017, and Wang *et. al.* in 2019, <sup>65,66,196</sup>. The most recent of these, from Wang and colleagues, used Illumina sequencing of the fungal ITS3-ITS4 region to identify exudate-utilising fungi associated with the rhizosphere of an unspecified variety of wheat. By using Illumina sequencing Wang and colleagues were able to survey the whole fungal community more precisely within the heavy and light fractions and were able to identify ten fungal groups which were able to utilise wheat root exudates.

The two other studies, from Ai *et al.*, and Uksa *et al.*, both focussed on the bacterial community. Comparing these two studies, the authors presented similar findings but with some distinct differences. Both studies showed that exudate-metabolising microbial communities in the rhizosphere consisted primarily of Actinobacteria and Proteobacteria <sup>65,196</sup>, though Uksa and colleagues also identified Bacteroidetes and Firmicutes as root exudate utilisers. Within the Proteobacteria Ai and colleagues showed Burkholderiales order taxa dominated exudate metabolism. Uksa and colleagues presented a higher taxonomic resolution despite using a shorter 330 base pair (bp) V1-V2 16S rRNA gene sequencing amplicon, which is unlikely to provide enough taxonomic resolution for genus-level identification <sup>132</sup>. In spite of this they proposed that a number of taxa were utilising root exudates including (among others) *Paenibacillus* and *Cohnella* within the Firmicutes, *Flavobacterium* and *Chitinophageaceae* within the Bacteroidetes, *Massilia*, *Variovorax*, and *Duganella* within the Proteobacteria, and *Kitasatospora*, *Promicromonospora*, and *Streptomyces* within the Actinobacteria <sup>196</sup>. Interestingly in this study *Streptomyces* were amongst the most highly abundant taxa within the wheat RAM at all sampling depths. The exudate utilising *Kitasatospora* genus are close relatives of *Streptomyces*, these groups can be difficult to distinguish from one another using 16S rRNA gene sequences, so it is unclear if this genus was truly utilising host

derived carbon within this study. This exemplifies the limits of short read amplicon sequencing.

The majority of exudate utilising Proteobacteria identified by Uksa and colleagues were from the Burkholderiales order, a similar observation to that of Ai and colleagues<sup>65,196</sup>. Whilst at the phylum level there were some similar findings from these two studies, there were also some key differences such as the identification of exudate utilising Bacteroidetes and Firmicutes by Uksa and colleagues. Given that these differences were observed at the phylum level this is unlikely to result from the differences in sequencing amplicon length between the two studies.

These contrasting observations could be due to the different varieties of wheat used in these studies, Ai and colleagues worked with *T. aestivum* cv. Shimai 18, whereas Ai and colleagues used the Scirocco variety. Potentially more importantly however is the fact that both studies used different soils. The soil used by Ai and colleagues was categorised as a sandy loam, sampled from an agricultural site in China and subject to fertilisation, whilst Uksa and colleagues sampled a mildly acidic silt loam soil from an arable field in Germany, and also sampled more alkaline clay silt loam from a greater depth to represent deeper soil. The growth and labelling conditions also varied between the studies. Whilst both were conducted in a greenhouse, they used different photoperiod lengths (12 or 16 hours), conducted <sup>13</sup>C CO<sub>2</sub> labelling at different growth timepoints (after 40 days compared to 75 days), and used different labelling period lengths (<sup>13</sup>CO<sub>2</sub> injection for eight hours per day over seven days, compared to 12 hours per day over 15 days). It is likely that all of these differences cumulatively resulted in the observed differences between the two studies, though differences in the soil type (for the reasons discussed in section 1.2.1) and labelling period are likely to have been the greatest contributors. For the labelling period, labelling of bacteria is dependent on incorporation of the <sup>13</sup>C label into the DNA backbone; this will only happen to a sufficient extent where a bacterium is actively growing and thus performing genome duplication<sup>183</sup>. During shorter labelling periods slower growing exudate utilising organisms, such as the *Streptomyces* identified by Uksa *et al.*<sup>196</sup>, may not become sufficiently labelled. The longer 15-day labelling period used by Uksa and colleagues may explain why they were able to identify a greater variety of exudate utilising taxa than Ai and colleagues<sup>65,196</sup>.

### 1.2.3.2 Limitations and considerations for the identification of root exudate utilising organisms using stable isotope probing

While SIP is a powerful technique for the study exudate utilising microorganisms, there are a number of considerations and caveats that must be accounted for when designing and conducting SIP experiments for the root microbial community.

Shorter labelling periods for example can underestimate the extent of exudate utilisation by slow growing and filamentous organisms such as *Streptomyces*; these filamentous bacteria expand primarily via hyphal tip extension, as such genome duplication does not occur as regularly when compared to bacteria which replicate via binary fission, therefore  $^{13}\text{C}$  will not be incorporated into the DNA backbone as quickly for this group. This means that short labelling periods can bias SIP studies towards faster growing organisms, representing a major caveat that must be considered when analysing data from DNA-SIP studies. To overcome this issue RNA SIP can be used, where RNA is isolated from samples instead of DNA. RNA synthesis occurs much more rapidly than DNA synthesis in active cells, and so by targeting RNA for analysis a much shorter  $^{13}\text{C}$  labelling period can be used <sup>183</sup>. For one study RNA was isolated just eight hours after introduction of the  $^{13}\text{C}$  labelled substrate, and the authors were able to successfully identify the microorganism degrading the substrate <sup>199</sup>. Whilst RNA is more difficult to isolate and less stable, this approach provides numerous advantages over DNA SIP for the rapid identification of active exudate utilising microorganisms. This approach has been used to identify exudate utilising organisms within the rhizosphere of wild plants *in situ* <sup>200</sup>, demonstrating the utility of short labelling periods.

Numerous factors can influence the outcome of SIP experiments probing plant-microbe interactions, as the interaction depends on root exudates many of the factors discussed in section 1.2.2.3 are important. Factors like plant species and developmental growth stage for example influence exudation rates <sup>73,143,144</sup>, and therefore the ability of root associated microorganisms to metabolise the exudates and incorporate the  $^{13}\text{C}$  label. In a given experiment these factors are important to control. The concentration of  $\text{CO}_2$  within the plant growth chamber also varies between studies <sup>65,66,161</sup>, though the impact of this on study outcomes is unclear. One other caveat that can be difficult to account for relates to the GC content of microbial genomes. DNA with a higher proportion of guanine and cytosine bases will be more dense, due to the higher molecular weight of these two nucleotides when compared to adenosine and thymine. As such, after density gradient ultracentrifugation microbial DNA with a higher GC content will naturally be present

within heavier fractions due to the higher density of high GC DNA. This makes it difficult to distinguish if high GC microorganisms are isotopically labelled, or if enrichment in heavy fractions is the result of a denser high GC genome<sup>183</sup>. To control for this <sup>12</sup>C unlabelled controls can be used. If a microorganism is enriched within the heavy fractions of the unlabelled <sup>12</sup>C control to the same or a greater extent than within the heavy fractions for the <sup>13</sup>C labelled plants, then it is likely that the presence of this microorganism within heavy fractions is the result of the GC content of the organism's genome and is not the result of root exudate utilisation. However, if a high GC microorganism is enriched to a greater extent within <sup>13</sup>C heavy fractions than within the <sup>12</sup>C heavy fractions, this would indicate <sup>13</sup>C incorporation and exudate utilisation<sup>183,201</sup>.

The presence of CO<sub>2</sub> fixing autotrophs must also be considered by studies using isotopically labelled CO<sub>2</sub>. Whilst some photoautotrophs reside within soils<sup>202,203</sup>, these organisms primarily reside within the photic zone, where photoautotrophy is possible<sup>202</sup>. Under the surface, where sampling from root associated regions occurs, no light can penetrate the soil, thus no photoautotrophy occurs. Many microbes residing within soils are chemoautotrophs however, for example *Pseudaminobacter salicylatoxidans*<sup>204</sup> or many ammonia oxidising archaea<sup>205</sup>, and so these microbes are able to assimilate atmospheric CO<sub>2</sub>. Within a SIP experiment this could result in non-exudate utilising microorganisms becoming isotopically labelled as these microbes can fix the <sup>13</sup>CO<sub>2</sub> from the chamber headspace and incorporate the <sup>13</sup>C label into their DNA independently of the host plant. To account for this, unplanted soil controls can be used; chambers containing unplanted soil are incubated with the <sup>13</sup>C labelled CO<sub>2</sub> and this community is also fractionated and profiled to identify taxa enriched within the heavy fractions, concluded to be CO<sub>2</sub> fixing autotrophs. These taxa can then be discounted from analysis of the root associated community, as their presence in the root associated heavy fractions is likely to be the result of chemoautotrophy, and not of exudate utilisation<sup>183,201</sup>.

SIP is a powerful tool that has been widely used to track the flow of nutrients and transfer of metabolites within microbial communities. Within the field of plant-microbe interactions SIP has proven a valuable tool for the identification of microorganisms interacting with the host via root exudates. While the method has numerous caveats, for example GC content or soil autotrophs, many of these can be controlled for within the experimental design. SIP indicates which microorganisms are able to utilise host-derived metabolites *in planta* and even *in situ* with wild plants, or plants grown in the field. The identification of these

organisms can preclude detailed studies to investigate which root exudates these organisms are consuming, how the plants exuded metabolites are able to modify the behaviour of these microorganisms and investigations of the mechanisms which underly the interaction between an exudate utilising organism and the host plant.

#### **1.2.4 Challenges and opportunities for studying the endosphere microbiome**

As has been alluded to in section 1.2.2.2., culture independent study of the microbiome within the endosphere compartment can be challenging due to contamination from host derived sequences. When DNA or RNA is extracted from the endosphere, the vast majority of the molecules recovered will be host derived, this means sequencing the microbial community to a sufficient degree of coverage to survey the diversity can be extremely challenging. For this reason, often PCR based amplicon sequencing methods are preferred, however host contamination can still cause significant issues. Universal primers used for amplification of the bacterial 16S rRNA gene, necessary to capture maximal bacterial diversity, will often also amplify host derived 16S rRNA gene sequences from the chloroplast or mitochondria. To mitigate this deeper sequencing must be used to ensure the microbial community is captured to sufficient depth. This in turn can also cause issues with the taxonomic resolution available for study as deep sequencing with longer read methods such as PacBio is highly expensive, and methods such as pyrosequencing are no longer commonly available. Depletion methods can also be used to reduce the quantity of host derived sequences, for example for human tissue microbiome studies saponin (a terpenoid compound) can be used to lyse host cells whilst leaving bacterial cells intact, so DNase treatment can be used to remove host DNA sequences prior to bacterial lysis <sup>206</sup>. A number of commercial kits also exist for host-DNA depletion, these work in a number of ways, for example depletion using column chromatography to bind and remove CpG methylated host DNA, or by differential lysis of host tissues. The efficiency of these kits however can vary <sup>207</sup>. There is indication that depletion methods may also have the potential to introduce bias, demonstrating that there are no perfect solutions to the challenges of host contamination <sup>207,208</sup>.

Another approach is to use peptide nucleic acid (PNA) blockers, these consist of a peptide conjugated to a primer specific to the host chloroplast or mitochondrial 16S rRNA gene sequence. The molecules are added to the PCR mixture prior to amplification where they anneal to the host-derived sequences, the conjugated

peptide will then block progression of the polymerase. In this way the extent of host sequence amplification can be greatly reduced, increasing the proportion of microbial 16S rRNA gene sequences within the amplicon and thus increasing microbial community coverage within the sequencing run <sup>209</sup>. Whilst this method does not completely eliminate host contamination, it can greatly reduce its extent, increasing the coverage for the microbial community within the sequencing run and making cheaper lower depth sequencing runs viable for microbiome analysis. For example, this can mean that longer read PacBio sequencing methods are viable for studying the endosphere community, as more affordable lower depth sequencing runs can yield usable quantities of data. Whilst PNA blockers are expensive to acquire, many commercial sequencing services offer to add them for a small fee, making the method broadly accessible.

The issue of host contamination is all the more difficult to overcome within metagenomic or metatranscriptomic studies; whilst host depletion methods can be used, without the use of PCR to selectively amplify microbial DNA sequences, sequencing to sufficient depth to gain good coverage for the microbiome is extremely difficult, at the time of writing no published study has achieved this for plant tissues.

### **1.2.5 The wheat seed microbiome and its influence on root microbial communities**

Whilst a number of studies have conclusively demonstrated that the vast majority of root associated microbiota are recruited from the bulk soil <sup>51–56</sup>, increasing attention is being paid to the role that the seed associated community may play in microbiome assembly. Some are now arguing that the role of the seed endophytic and epiphytic community has been understudied, and may be significant <sup>202,210,211</sup>. Many bacteria and fungi have been isolated from seeds <sup>212,213</sup>, and a number of these have been shown to possess plant beneficial traits <sup>214–216</sup>. Some biotechnological applications are also being developed to inoculate the seed with beneficial microorganisms as a delivery mechanism for biocontrol strains <sup>217</sup>. For barley beneficial taxa such as *Paenibacillus*, *Pseudomonas*, and *Pantoea* were identified within the seeds, and isolates from these groups were able to relieve abiotic stress and to colonise the rhizosphere <sup>216</sup>.

Whilst these beneficial microbes are present within the seeds, their prevalence within the seed associated microbiome is uncertain as this type of strain-level

analysis is difficult to conduct from plant tissues using high-throughput DNA sequencing methods, as discussed in section 1.2.4. For similar reasons it also remains unclear whether these organisms actually colonise plant associated niches from the seed, or if they are outcompeted by equivalent soil microorganisms. To answer these questions strain-level analysis from the root and seed endosphere compartments would be required. Some studies have claimed to have shown that specific archaeal and bacterial lineages are transmitted between generations of plants via the seeds <sup>218–221</sup>; short-read amplicon sequencing however does not provide sufficient taxonomic resolution to draw such conclusions, these studies cannot account for changes in the abundance of individual strains that are masked by the family level taxonomic assignments provided by short read amplicon sequencing. This means that such studies cannot eliminate the possibility that the strains which colonised the seeds are different to those which are colonising the roots. It is also impossible to rule out the possibility that the same taxa had colonised the roots from the soil. In the future culture-based experiments are required to answer this question, where known strains can be synthetically labelled and either fluorescence microscopy or qPCR with synthetic barcodes can be used to probe if or how these strains migrate from the seeds to the roots, and the prevalence of these organisms within the seed can be compared to that within the root.

Despite the challenges posed by studying the seed microbial community, a number of studies have attempted to characterise the seed microbiome of wheat. For the epiphytic community for example (on the seed surface) the community was found to be remarkably similar when comparing wheat (bread and durum) to a range of *Brassica* species (brown mustard (*Brassica juncea* L. Czern), field mustard (*Brassica rapa*), and oilseed rape (*Brassica napus*)). A number of bacterial isolates from these seeds were also able to impede the growth of fungal isolates <sup>213</sup>. Further to this, plant growth promoting bacteria have been isolated from the wheat seed endosphere <sup>215</sup>, indicating that agriculturally useful microbiota may reside within the seed. The fungal and bacterial communities within wheat seeds have been shown to vary significantly by genotype, and when cultivated axenically more diverse communities colonise the roots from wild varieties when compared to domesticated wheat <sup>129</sup>. The bacterial and fungal epiphytic community is also influenced by genotype, and by the environmental conditions within the field <sup>222</sup>. In spite of this, recent work indicated that the seed microbiome had no effect on the rhizosphere community of wheat <sup>136</sup>. In context with the broader literature this would lead us to

hypothesise that beneficial microbiota may colonise the seed as a result of successful root colonisation from the soil, but that they are unable to colonise the root microbiome as the seed germinates and are outcompeted by soil microorganisms. It remains unknown however how microorganisms colonise the seed endosphere, or the extent to which beneficial seed endosphere microorganisms can colonise the plant. Given the presence of beneficial microorganisms within the seeds, more investigations are required to further characterise microbial diversity within the seeds, to understand how the seeds are colonised and what effect, if any, the seed microbiome may have on wheat.

### **1.2.6 Developmental senescence and its potential influence on plant root microbiomes**

For many important crops such as wheat, barley, maize, corn, and rice, developmental senescence is a crucial determinant of yield and nutrient content<sup>223,224</sup>. Developmental senescence (often visible by the characteristic yellowing of crops in late summer) occurs at the end of the growing season, and during this process the plants resources, particularly nitrogen, are diverted from tissues into the developing grain<sup>223,224</sup>. Senescence represents a dramatic shift in the metabolic activity of the plant<sup>223</sup> and in the regulation of numerous pathways for pathogen defence<sup>224,225</sup>. Given that root exudation is a highly dynamic process<sup>140</sup>, it would be reasonable to assume that senescence affects root exudation substantially, particularly because of the diversion of nitrogen to the developing grain. Several major wheat root exudate compounds, such as amino acids, nucleosides, and organic acids, all contain nitrogen<sup>140</sup>. The production and exudation of these compounds is therefore likely to be altered or attenuated during developmental senescence. Given the pivotal role senescence plays in grain development and yields, microbial community dynamics during this process warrant investigation.

At the time of writing no studies have specifically investigated microbiome dynamics as plants senesce. At the onset of senescence, plant resources are redirected to the seed, root exudation is reduced, and root tissues start to decay. It is plausible that this shift in plant metabolism would cause a change in the root-associated microbiome, and so greater understanding of this shift could come inform agricultural management strategies and the design of new crop cultivars. It could also be indicative of which root associated microbiota require active input from the

living plant to persist within the root associated microbiome, providing further clues as to which microbiota are core members of the root microbiome.

### 1.2.7 Archaea within plant root microbiomes

Archaea are a hugely diverse group of microorganisms present in a broad variety of environments such as intertidal mangrove sediments<sup>226</sup>, rice paddies<sup>227,228</sup>, marine sponges<sup>229</sup>, human skin<sup>230</sup>, and freshwater river sediments<sup>231</sup>. Archaea can also be found in many soils<sup>232</sup>, including agricultural soils<sup>205,233,234</sup>. Most generic and commonly used 16S rRNA gene PCR primer sets fail to capture archaeal diversity<sup>235</sup>, despite the assumption by many studies that these primers are sufficient to capture archaeal diversity. The diversity of archaea within agricultural soils and plant associated niches is therefore commonly overlooked. Key soil groups such as ammonia oxidizing archaea (AOA) play a significant role in nitrogen cycling, a key ecological service, and one study has managed to link an AOA to plant beneficial traits<sup>236</sup>, suggesting that the role of archaea within the root associated microbiome warrants further study.

Within mangrove ecosystems the anoxic sediment is home to methanogens, ammonia oxidisers, and sulphate reducing archaea<sup>226,237</sup>. Within the rhizosphere of different species of mangrove trees, distinct archaeal communities can be found, and these archaea are involved in nutrient cycling through methanogenesis and sulphate reducing activity<sup>238</sup>. It is unclear whether methane production is influenced by the host plant. Other wetland habitats also host archaea within plant associated niches, most commonly methanogens<sup>92,239-241</sup>. There is evidence that for some wetland plant species (but not all) the abundance and activity of AOA is reduced within the rhizosphere<sup>241</sup>, demonstrating the potential for a highly diverse range of niches for archaea within botanically rich environments.

Rice paddies, another anoxic and waterlogged environment, have been studied extensively for the role of archaea within plant associated niches, and in particular methanogens. A number of studies have demonstrated that methanogens and AOA reside within the rhizosphere of rice plants<sup>228,242,243</sup>, and that microbial diversity is influenced by soil type, host genotype, and crop management practices<sup>50</sup>. Plant roots are an oxygenated environment. As methanogens typically reside within anoxic sediments it remains unclear how they have adapted to thrive within this niche. There is some literature suggesting that these archaea utilise host-derived carbon, and thus that methane production from these organisms is dependent on

photosynthetically fixed carbon from the plant. This conclusion has been drawn as when rice was incubated with  $^{13}\text{CO}_2$  during the photoperiod,  $^{13}\text{C}$ -methane ( $^{13}\text{CH}_4$ ) could be detected in the headspace, and methanogen 16S rRNA gene sequences became isotopically labelled <sup>191,242</sup>. This would provide key evidence demonstrating the capability of rhizosphere archaea to interact with the host and utilise host derived metabolites. Many methanogens however, for example those using either the hydrogenotrophic or methylotrophic methanogenesis pathways, are able to fix  $\text{CO}_2$  autotrophically, and then use this carbon to produce methane <sup>244</sup>. Whilst the  $\text{CO}_2$  autotrophy-dependant hydrogenotrophic pathway is the most widespread of these <sup>244</sup>, methanogens have often been documented using alternate substrates such as acetate (acetoclastic methanogeny), formate, or various methylated compounds <sup>245</sup>, so it is possible that in these studies the isotopically labelled methanogens are utilising host derived carbon. Given however that these studies did not provide unplanted controls to account for  $\text{CO}_2$  autotrophy, it remains unclear if methanogens residing within the rhizosphere do indeed utilise host exudates. This represents a significant gap in our knowledge surrounding how archaea residing within plant associated habitats might interact with the host plant.

While much work has been done to characterise archaea-plant interactions within anoxic waterlogged ecosystems, archaea are prevalent within terrestrial agricultural ecosystems. Soil archaeal diversity is comprised of AOA primarily <sup>67,246–248</sup>. It is well documented that AOA play a major role in the cycling of nitrogen within agricultural soils as they are responsible for the oxidation of ammonium, which is added to agricultural ecosystems as fertiliser <sup>247,248</sup>. Typically, AOA are associated with the loss of bioavailable nitrogen from agricultural systems as they are responsible for the oxidation of ammonium to nitrate, which is then leached from the soil and can cause ecological problems such as hypoxia and eutrophication within waterways <sup>247</sup>. A number of studies have investigated the influence a number of agricultural practices have on the archaeal soil community. The abundance of AOA has been shown to increase within the rhizosphere of plants cultivated in agricultural soil for example <sup>247</sup>, and a number of management practices such as tillage and fertilisation have been shown to effect archaeal community diversity <sup>67,246,249</sup>. The presence of genes encoding for archaeal siderophore production has also been correlated with disease suppression within disease suppressive agricultural soils <sup>250</sup>. These studies demonstrate that archaea are a significant component of agricultural ecosystems, and that further investigation of their role within the root microbiome of crops is warranted.

Increasingly studies are aiming to identify archaeal diversity within the roots of terrestrial plants. The majority of these studies identified phyla such as *Crenarchaeota* or *Euryarchaeota* (Table 1.2), high-level taxonomic identifications which make the function of these microorganisms difficult to postulate. Compared to bacterial resources, archaeal databases are far less comprehensive and until recently lacked the established framework of their bacterial counterpart<sup>251</sup>. Recently however a comprehensive phylogenetic framework for archaeal diversity was published for the genome taxonomy database (GTDB<sup>252</sup>). Most phyla within the *Crenarchaeota* have now been recategorized into the Thermoproteota phylum. This group still contains a large amount of archaeal diversity, including the Nitrososphaerales order (ammonia oxidising archaea) and the Methanomethylales order (methanogens), making the identity and function of historically identified *Crenarchaeota* difficult to postulate (though in most soil systems they are most likely to be represented by AOA). The *Euryarchaeota* phylum has been split into several phyla containing a range of archaeal diversity, such as methanogens and halophiles, making the function of historical *Euryarchaeota* identifications even more difficult to postulate. Though this reorganisation has made high level taxonomic identifications in older studies more difficult to interpret, this newly established framework will aid future studies using this database to survey archaeal diversity within plant associated niches. Whilst this represents excellent progress, a large proportion of archaeal diversity remains uncultured and uncharacterised, with can still make the ecological role of uncultivated archaeal taxa detected by high-throughput sequencing experiments elusive.

Table 1.2. Studies identifying interactions or associations between archaea and terrestrial plant associated niches

Plant species & environment	Archaeal taxa identified	Observed interaction	References
<b><i>Arabidopsis thaliana</i></b> Grown under sterile conditions or in potting compost	<i>Nitrosocosmicus oleophilus</i> MY3	An ammonia oxidising archaeon can promote the growth of <i>A. thaliana</i> and provide protection from bacterial pathogens via induced systemic resistance	Song <i>et al.</i> (2019) <sup>236</sup>
Bryophyte moss, Lycopod club moss. Pteridophyte fern, Gymnosperm conifer, Dicot, and Monocot, plant rhizospheres sampled from a range of soils	<i>Crenarchaeota</i>	PCR-SSCP electropherograms showed evidence for the rhizosphere effect acting upon the archaeal community	Sliwinski & Goodman (2004) <sup>255</sup>
<b>Coffee (<i>Coffea arabica</i>)</b> Coffee cherry endosphere, sampled from coffee plantations in Brazil, acidic soils	<i>Crenarchaeota</i> , <i>Halocooccus</i> , <i>Haloferax</i> , <i>Haobacterium</i> , <i>Thermoplasma</i>	Archaeal 16S rRNA gene sequences were recovered from the interior of coffee cherries, it was hypothesised that halophilic archaea are adapted to live within the osmotically stressful coffee cherries	Oliveira <i>et al.</i> (2013) <sup>253</sup>
<b>Tree species</b> <i>Pinus sylvestris</i> & <i>Picea abies</i> , & deciduous trees ( <i>Betula pendula</i> & <i>Alnus glutinosa</i> ) rhizosphere, grown in pine forest soil	<i>Crenarchaeota</i> , <i>Halobacteriales</i>	Different tree species harbour distinct archaeal rhizosphere communities, and ectomycorrhizal colonisation correlated with an increased archaeal load and the presence of 1.1c <i>Crenarchaeota</i> .	Bomberg & Timonen (2009) <sup>256</sup>
Onderosa Pine, Sitka Spruce, and Western Hemlock rhizospheres, grown in forest soil	n/a	qPCR showed a reduction in the abundance of archaea within the rhizosphere and mycosphere, and correlated increased exudation with reduced archaeal abundance	Karlsson <i>et al.</i> (2011) <sup>168</sup>
<i>Suillus bovinus</i> & <i>Pinus sylvestris</i> rhizospheres and root tissues, sampled from a pine forest	Primarily <i>Thaumarchaeota</i>	Archaeal abundance was greater within fine short roots and mycorrhizal tissue, and different <i>Thaumarchaeotal</i> lineages were detected in different root compartments	Rinta-Kanto & Timonen <i>et al.</i> (2020) <sup>257</sup>
<i>Picea crassifolia</i> & <i>Populus szechuanica</i> rhizospheres sampled from a semi-arid region of the Qinghai-Tibetan plateau	Primarily <i>Thaumarchaeota</i> , <i>Euryarchaeota</i> , <i>Bathyarchaeota</i>	Rhizosphere archaeal community differed significantly by soil type and plant species, and an unclassified archaeal OTU may be responsible for archaeal species interactions	Zhang <i>et al.</i> (2020) <sup>258</sup>

<p><b><i>Jatropha curcas</i> (a biodiesel crop)</b></p> <p>Rhizosphere, grown a range of neutral to alkaline pH agricultural soils</p>	<p>Primarily <i>Crenarchaeota</i>, <i>Euryarchaeota</i>, <i>Methanomicrobiaceae</i></p>	<p>A diverse community of archaea were found to reside within the rhizosphere, and this varied significantly by soil type</p>	<p>Dubey <i>et al.</i> (2016) <sup>259</sup></p>
<p><b>Maize (<i>Zea mays</i>)</b></p> <p>Endosphere, grown in acidic agricultural soil</p> <p>Endosphere and rhizosphere, grown in acidic agricultural soil</p> <p>Rhizosphere cultivated in agricultural soil</p>	<p><i>Crenarchaeota</i>, <i>Euryarchaeota</i></p> <p><i>Crenarchaeota</i>, <i>Euryarchaeota</i>, <i>Thaumarchaeota</i></p> <p>Ammonia oxidising archaea</p>	<p>Archaeal groups were present within 16S rRNA gene clone libraries.</p> <p>Application of organic fertiliser was associated with increased archaeal diversity within the rhizosphere and endosphere</p> <p>The abundance of AOA was found to increase within the rhizosphere</p>	<p>Chelius &amp; Triplett (2001) <sup>260</sup></p> <p>Fadji <i>et al.</i> (2020) <sup>261</sup></p> <p>Wattenburger <i>et al.</i> (2020) <sup>247</sup></p>
<p><b>Mediterranean olive trees (<i>Olea europea</i>)</b></p> <p>Phyllosphere, sampled from orchards around the Mediterranean</p>	<p>Primarily <i>Methanomicrobiales</i>, <i>Halobacteriales</i>, <i>Nitrososphaerales</i>, <i>Crenarchaeota</i></p>	<p>A diverse community of archaea was found to reside in olive branches and leaves, and this community varied significantly by location</p>	<p>Müller <i>et al.</i> (2015) <sup>253</sup></p>
<p><b>Rocket lettuce (<i>Eruca sativa</i>)</b></p> <p>Phyllosphere and rhizosphere, grow in suburban garden soil</p>	<p>Primarily <i>Nitrosocosmicus</i>, <i>Thaumarchaeota</i>, <i>Methanosarcina</i></p>	<p>The phyllosphere and rhizosphere both harboured a distinct archaeal community compared to the soil</p>	<p>Taffner <i>et al.</i> (2019) <sup>262</sup></p>
<p><b>Switchgrass (<i>Panicum virgatum</i>) and Miscanthus (<i>Miscanthus x giganteus</i>)</b></p> <p>Grown in agricultural soil</p>	<p>n/a</p>	<p>The archaeal community within the phyllosphere of two biofuel crops varied significantly by season</p>	<p>Grady <i>et al.</i> (2019) <sup>62</sup></p>
<p><b>Tomato (<i>Solanum lycopersicum</i>)</b></p> <p>Grown in potting compost</p> <p>Multiple cultivars grown in agricultural soils</p>	<p><i>Crenarchaeota</i></p> <p><i>Thaumarchaeota</i>, <i>Methanosarcina</i>, <i>Methanoculleus</i></p>	<p>Filtered root tissue extract was able to enrich for archaea from agricultural soil within an enrichment culture experiment</p> <p>The quantity of archaea within the rhizosphere and soil was similar, and archaea were also detected within the seeds. Archaeal community composition was influenced genotype.</p>	<p>Simon <i>et al.</i> (2005) <sup>263</sup></p> <p>Taffner <i>et al.</i> (2020) <sup>219</sup></p>
<p><b>Wheat</b></p> <p><i>Triticum aestivum</i>, multiple genotypes grown in a range of agricultural soils</p> <p><i>Triticum turgidum</i>, multiple genotypes grown in a soil sand mixture</p>	<p><i>Nitrososphaeraceae</i>, <i>Nitrosotaleaceae</i></p> <p>n/a</p>	<p>Ammonia oxidising archaea were identified associated with the wheat rhizosphere across multiple soil types</p> <p>The diversity of archaea varied amongst the different wheat cultivars</p>	<p>Simonin <i>et al.</i> (2020) <sup>128</sup></p> <p>Iannucci <i>et al.</i> (2021) <sup>114</sup></p>

A number of studies have managed to identify and characterise archaea-plant interactions beyond the phyla level. Oliveira and colleagues for example managed to identify a number of archaeal families within the endosphere of coffee cherries (Table 1.2). The identification of halophilic families of archaea in this environment, typically found in extremely salty environments, lead the authors to hypothesise the presence of halophile-like archaea, adapted to live within the osmotically stressful environment generated within coffee cherries by high oligosaccharide concentrations <sup>253</sup>. Other studies were also able to identify families or genera of AOA such as *Nitrosocosmicus* or *Nitrosostaleaceae* <sup>128,219,254</sup>, allowing the authors to conclude with relative confidence that a AOA are present within the root associated niche. Most notably there has been one culture-based study investigating the interaction between an ammonia oxidising archaeon and *A. thaliana*. This is the first example of a direct investigation of the capacity of an archaeon to effect plant growth and health, and the authors were able to show that the AOA strain *Nitrosocosmicus oleophilus* MY3 can both promote the growth of the host plant, and provided protection from disease via induced systemic resistance (ISR) <sup>236</sup>. This intriguing result indicates a greater importance for archaea-plant interactions than had been previously postulated and warrants further culture-based experiments investigating the interaction between archaea and terrestrial plants.

Overall, archaea are a large and diverse group of organisms found in a range of habitats. Examples from wetland ecosystems and rice paddies demonstrate the potential for interactions between archaea and plants, and their prevalence in agricultural soil further supports this potential. In recent years a new interest in the potential of archaea-plant interaction has spurred a number of studies to investigate these interactions. Despite this, in many settings the influence of archaea is still understudied, and our ability to study these organisms is limited. More research into the diversity of archaea within plant associated habitats, and into the interactions that occur between archaea and plants, is needed to fully understand the niche occupied by these organisms within agricultural plant microbiomes.

### 1.3 Project aims

This project aimed to help to address knowledge gaps pertaining to the core wheat root microbiome in a number of ways. Firstly, by helping to define the core root associated taxa, to attempt to identify these organisms to the genus/species level, and to investigate the influence of host genotype, soil type, and of developmental senescence on community composition. This project also aimed to investigate the role of ammonia oxidising archaea within the root, and to characterise the wheat seed endophytic microbiome. To do so this project aimed to address the following questions-

1. Can we detect any core bacterial, fungal, or archaeal taxa associated with wheat?
2. How do key factors such as host genotype, soil type, and developmental senescence effect archaeal, bacterial, and fungal microbiome composition?
3. Which of these core microbes, if any, are able to utilise host derived carbon within root associated niches?
4. What role do ammonia oxidising archaea play within the root?

## Chapter 2. Materials and methods

Some material from this chapter has been published previously <sup>198</sup>.

**Table 2.1. Media Recipes**

Medium	Uses	Recipe
Lysogeny Broth (LB)	General growth and maintenance of strains, bioassays	<ul style="list-style-type: none"> <li>- 5 g Yeast Extract</li> <li>- 10 g NaCl</li> <li>- 10 g Tryptone</li> <li>- Up to 1 L dH<sub>2</sub>O</li> </ul> For solid medium 2% (vol/vol) agar was used
Reasoners 2A agar (R2A)	Cultivation of endophytes	<ul style="list-style-type: none"> <li>- 0.5 g Yeast Extract</li> <li>- 0.5 g Protease Peptone</li> <li>- 0.5 g Casamino acids</li> <li>- 0.5 g Glucose</li> <li>- 0.5 g Soluble Starch</li> <li>- 0.3 g Sodium Pyruvate</li> <li>- 0.3 g K<sub>2</sub>HPO<sub>4</sub></li> <li>- 0.05 g MgSO<sub>4</sub></li> <li>- Up to 1 L dH<sub>2</sub>O</li> </ul> - For solid medium 1.5% (vol/vol) agar was used - Where oxalic acid was used (R2A-O), 4.5 g was added per litre of medium (50 mM) - Medium was pH adjusted to 7.72 with NaOH
BAz	Cultivation of endophytes	<ul style="list-style-type: none"> <li>- 2 g Azelaic Acid</li> <li>- 0.2 g L-Citrulline</li> <li>- 0.5 g K<sub>2</sub>HPO<sub>4</sub></li> <li>- 0.2 g MgSO<sub>4</sub></li> <li>- Up to 1 L dH<sub>2</sub>O</li> </ul> - For solid medium 1.5% (vol/vol) agar was used - Where oxalic acid was used (BAz-O), 4.5 g was added per litre of medium (50 mM) - Medium was pH adjusted to 5.72 with NaOH
MAG	Cultivation of endophytes	<ul style="list-style-type: none"> <li>- 1.1 g MES Sodium Salt</li> <li>- 1.3 g HEPES</li> <li>- 0.007 g FeCl<sub>2</sub></li> <li>- 0.015 g CaCl<sub>2</sub></li> <li>- 1 g Yeast Extract</li> <li>- 1 g D-Gluconic Acid</li> <li>- 1 g L-Arabinose</li> <li>- 0.22 g K<sub>2</sub>HPO<sub>4</sub></li> <li>- 0.18 g MgSO<sub>4</sub></li> <li>- 0.25 g NaSO<sub>4</sub></li> <li>- Up to 1 L dH<sub>2</sub>O</li> </ul> - For solid medium 1.5% (vol/vol) agar was used - Medium was pH adjusted to 6.45

<i>Burkholderia cepacia</i> selective agar (BCSA)	Cultivation of endophytes	<ul style="list-style-type: none"> <li>- 10 g Casein Peptone</li> <li>- 10 g Lactose</li> <li>- 10 g Sucrose</li> <li>- 5 g NaCl</li> <li>- 1.5 g Yeast Extract</li> <li>- Up to 1 L dH<sub>2</sub>O</li> </ul> <p>- For solid medium 1.5% (vol/vol) agar was used</p>
Nutrient Agar (NA)	Cultivation of endophytes	<ul style="list-style-type: none"> <li>- 4 g Difco Nutrient Broth</li> <li>- Up to 1 L dH<sub>2</sub>O</li> </ul> <p>For solid medium 1% (vol/vol) agar was used</p>
<i>Nitrosocosmicus</i> / ammonia oxidising archaea (AOA) medium	Cultivation of ammonia oxidising archaea	<p><u>10x salts solution</u></p> <ul style="list-style-type: none"> <li>- 10 g NaCl</li> <li>- 4 g MgCl<sub>2</sub></li> <li>- 1 g CaCl<sub>2</sub></li> <li>- 2 g KH<sub>2</sub>PO<sub>4</sub></li> <li>- 5 g KCl</li> </ul> <p>After autoclaving, the 10x salts solution was diluted to 1x in sterile dH<sub>2</sub>O, to total volume of 800 ml. Then the following additives are added</p> <ul style="list-style-type: none"> <li>- 0.8 ml Modified Trace Elements (Table 2.2)</li> <li>- 0.8 ml FeNaETDA Solution (Table 2.2)</li> <li>- 1.6 ml 1 M Sodium Bicarbonate</li> <li>- 4 ml 1 M NH<sub>4</sub>Cl</li> <li>- 8 ml HEPES buffer (Table 2.2)</li> <li>- 0.8 ml Vitamin Solution (Table 2.2)</li> <li>- 0.8 ml Phenol Red Solution (Table 2.2)</li> </ul> <p>All <i>Nitrosocosmicus</i> medium was made in acid washed glassware</p>
Murashige and Skoog (MSK)	In vitro plant cultivation and AOA PGP experiments	<ul style="list-style-type: none"> <li>- 4.43 g MSK Salts Medium</li> <li>- Up to 1 L dH<sub>2</sub>O</li> </ul> <p>For half strength MSK 2.215 g were used in 1 L</p>
Potato Dextrose Agar (PDA)	Take-all fungus cultivation and bioassays	<ul style="list-style-type: none"> <li>- 3.9 g Sigma Potato Dextrose Agar</li> <li>- Up to 1 L dH<sub>2</sub>O</li> </ul>
Water Agar (WA)	Sterile wheat cultivation	<ul style="list-style-type: none"> <li>- 2 g agar</li> <li>- Up to 100 ml dH<sub>2</sub>O</li> </ul>
Minimal medium (MM)	<i>Streptomyces</i> interaction assay	<ul style="list-style-type: none"> <li>- 0.5 g L-asparagine</li> <li>- 0.5 g K<sub>2</sub>HPO<sub>4</sub></li> <li>- 0.2 g MgSO<sub>4</sub></li> <li>- 0.01 g FeSO<sub>4</sub></li> <li>- 10 g Glucose (added after autoclaving)</li> <li>- Up to 1 L dH<sub>2</sub>O</li> </ul> <p>1% agar</p>

**Table 2.2 Buffer recipes**

Buffer	Uses	Recipe
Phosphate Buffered Saline (PBS)	Root sampling	<ul style="list-style-type: none"> <li>- 6.33 g NaH<sub>2</sub>PO<sub>4</sub></li> <li>- 16.5 g Na<sub>2</sub>HPO<sub>4</sub></li> <li>- 1 L dH<sub>2</sub>O</li> <li>- 0.02% Silwett L-77 (v/v)</li> </ul>
Tris Borate EDTA Buffer (TBE)	Agarose gel electrophoresis	<ul style="list-style-type: none"> <li>- 90 mM Tris HCl</li> <li>- 90 mM Boric Acid</li> <li>- 2 mM EDTA</li> </ul>
5x Bromophenol Blue Loading Buffer (5x)	DNA loading for agarose gel electrophoresis	<ul style="list-style-type: none"> <li>- 0.25% (w/v) bromophenol blue</li> <li>- 0.25% (w/v) xylene-cyanol blue</li> <li>- 40% (w/v) sucrose in water</li> </ul>
Tris Acetate EDTA Buffer (TAE)	Denaturing gradient gel electrophoresis	<ul style="list-style-type: none"> <li>- 242 g Tris-base</li> <li>- 57.1 ml acetic acid</li> <li>- 100 ml 0.5M EDTA</li> </ul> <p>The buffer was pH adjusted to 8</p>
Gradient Buffer	Density gradient ultracentrifugation	<ul style="list-style-type: none"> <li>- 0.1 M Tris-HCl (pH adjusted to 8)</li> <li>- 0.1 M KCl</li> <li>- 1 mM EDTA</li> </ul>
PEG-NaCl solution	Density gradient ultracentrifugation, DNA precipitation	<ul style="list-style-type: none"> <li>- 30% w/v polyethylene glycol 6000</li> <li>- 1.6 M NaCl</li> </ul>
HEPES Buffer	<i>Nitrosocosmicus</i> medium	<ul style="list-style-type: none"> <li>- 1 M HEPES</li> <li>- 0.6 M NaOH</li> </ul>
FeNaETDA Solution	<i>Nitrosocosmicus</i> medium	<ul style="list-style-type: none"> <li>- 7.5 mM FeNaETDA in dH<sub>2</sub>O</li> </ul> <p>Solution was filter sterilised and stored in the dark at 4°C</p>
Modified Trace Elements	<i>Nitrosocosmicus</i> medium	<ul style="list-style-type: none"> <li>- 100 mM HCl (~12.5M)</li> <li>- 0.5 mM H<sub>3</sub>BO<sub>3</sub></li> <li>- 0.5 mM MnCl<sub>2</sub></li> <li>- 0.8 mM CoCl<sub>2</sub></li> <li>- 0.1 mM NiCl<sub>2</sub></li> <li>- 0.01 mM CuCl<sub>2</sub></li> <li>- 0.5 mM ZnSO<sub>4</sub></li> <li>- 0.15 mM Na<sub>2</sub>MoO<sub>4</sub></li> </ul> <p>The solution was autoclaved and stored in the dark at 4°C</p>
Vitamin Solution	<i>Nitrosocosmicus</i> medium	<ul style="list-style-type: none"> <li>- 0.05 g Biotin</li> <li>- 0.02 g Folic Acid</li> </ul>

		<ul style="list-style-type: none"> <li>- 0.1 g Pyridoxine HCl</li> <li>- 0.05 g Thiamine HCl</li> <li>- 0.05 g Riboflavin</li> <li>- 0.05 g Nicotinic Acid</li> <li>- 0.05 g DL Pantothenic Acid</li> <li>- 0.05 g P Aminobenzoic Acid</li> <li>- 2 g Choline Chloride</li> <li>- 0.01 g Vitamin B12</li> <li>- Up to 1 L dH<sub>2</sub>O</li> </ul> <p>Solution was pH adjusted to 7 with KOH, filter sterilised and stored at 4°C</p>
Phenol Red Solution	<i>Nitrosocosmicus</i> medium	<ul style="list-style-type: none"> <li>- 0.05 g Phenol Red</li> <li>- Up to 100 ml dH<sub>2</sub>O</li> </ul> <p>Solution was filter sterilised and stored at 4°C</p>

## 2.1 Chemicals and reagents

Analytical grade reagents used in this work were purchased from Sigma-Aldrich (MIS, USA) or Fisher Scientific (Loughborough, UK). Molecular biology grade reagents were purchased from Sigma-Aldrich (MI, USA), Fisher Scientific (Loughborough, UK), Thermo Fisher (MA, USA), New England Biolabs (MA, USA), Quiagen (Germany), Promega UK (Southampton, UK), Cambridge Bioscience (Cambridge, UK) or Roche (Switzerland). Gasses were supplied by either BOC (UK) or Cambridge Isotope Laboratories (MA, USA).

## 2.2 Environmental sampling

### 2.2.1 Agricultural soil sampling and chemical analysis

Agricultural soil was sampled in April 2019 from the John Innes Centre (JIC) Church Farm cereal crop research station in Bawburgh, Norfolk, United Kingdom (52°37'39.4"N 1°10'42.2"E). The top 20 cm of soil was removed prior to sampling. For bulk soil associated with field grown wheat plants, sampling was performed in the same way, from bare soil approximately 30 cm away from the plant. Soil was stored at 4°C and pre-homogenised prior to use for cultivation, or snap frozen in liquid nitrogen and stored at -80°C for DNA extractions.

Chemical analysis was performed by the James Hutton Institute Soil Analysis Service (Aberdeen, UK) to measure soil pH, organic matter (%), and the

phosphorus, potassium, and magnesium content (mg/kg) (Table 2.3). To quantify inorganic nitrate and ammonium concentrations a KCl extraction was performed, where 3 g of each soil type suspended in 24 ml of 1 M KCl in triplicate and incubated for 30 minutes with shaking at 250 rpm. To quantify ammonium concentration (g/kg) the colorimetric indophenol blue method was used<sup>264</sup>. For nitrate concentration (g/kg) vanadium (III) chloride reduction coupled to the colorimetric Griess reaction as previously described in Miranda *et al.*<sup>265</sup>. The agricultural soil was mildly alkaline (pH 7.97), contained only 2.3% organic matter and was relatively low in inorganic nitrogen, magnesium, and potassium. Levington F2 compost was acidic (pH 4.98) and had a high organic matter content (91.1%) as well as higher levels of inorganic nitrogen, phosphorus, potassium, and magnesium (Table 2.3).

**Table 2.3. Soil chemical properties**

Parameter	Measurement	SAC Rating
<b>Agricultural Soil, John Innes Centre Field Studies Site (sampled 18.04.2019)</b>		
pH	7.97	n/a
Phosphorus (mg/kg)	81.63	High
Potassium (mg/kg)	103	Moderate
Magnesium (mg/kg)	34.8	Very Low
Nitrate (g/kg)	0.55	n/a
Ammonium (g/kg)	0.0035	n/a
Organic matter (%)	2.26	n/a
<b>F2 Levington Compost</b>		
pH	4.98	n/a
Phosphorus (mg/kg)	880.5	Excessively high
Potassium (mg/kg)	2508	Excessively high
Magnesium (mg/kg)	6021	Excessively high
Nitrate (g/kg)	4.36	n/a
Ammonium (g/kg)	0.1	n/a
Organic matter (%)	91.08	n/a

### 2.2.2 Field wheat sampling

Triplicate Paragon var. *Triticum aestivum* plants were sampled during the stem elongation growth phase approximately 200 days after sowing, in July 2019. To assess microbial diversity after senescence, triplicate Paragon var. *T. aestivum* plants were sampled immediately before harvest in August 2020 approximately 230 days after sowing. All field grown plants were sampled from the JIC Church Farm field studies site in Bawburgh (Norfolk, United Kingdom) (52°37'42.0"N 1°10'36.3"E)

and were cultivated in the same field from which agricultural soil was sampled. Plants were processed as described in 2.3.4.1.

## **2.3 Plant cultivation, sampling, and experimental methods**

### **2.3.1 Wheat seed surface sterilisation & cultivation**

All *T. aestivum* seeds were soaked for two minutes in 70% ethanol (v/v), 10 minutes in 3% sodium hypochlorite (v/v) and washed 10 times with sterile water to sterilise the seed surface. Seeds were then sown into pots of pre-homogenised, pre-wetted Church farm agricultural soil, Levington F2 compost, or a 50:50 (v/v) mix of the two. For all metabarcoding, plant growth promotion, or endophyte isolation experiments with pot cultivated plants, plants were propagated for 30 days prior to sampling. For stable isotope probing this period was 42 days, and all plant cultivation was at 21°C under a 12 h light/ 12 h dark photoperiod.

### **2.3.2 Sterile wheat seedling sampling**

*T. aestivum* var. Paragon seeds were surface sterilised as described in 2.3.1. Seeds were then placed on water agar (Table 2.1) and incubated at 4°C for three days to synchronize germination. Then plants were incubated at 21°C over four days until roots and shoots had begun to emerge. Under a Bunsen flame a sterile scalpel was used to remove the roots, taking care not to rupture the seed endosperm. Two plants were sampled and pooled per replicate, and a total of three such replicates were sampled and snap frozen in liquid nitrogen. Pooled root samples were then crushed in liquid nitrogen using a pestle and mortar under laminar flow before DNA was extracted using the FastDNA™ SPIN Kit for Soil (MP Biomedical) according to manufacturer's protocol with minor modifications: incubation in DNA matrix buffer was performed for 12 minutes and elution carried out using 75 µl DNase/Pyrogen-Free Water. All DNA samples were stored at -20°C and quality and yields were then assessed using a nanodrop and Qubit fluorimeter. DNA samples were then diluted to 50 ng / µl, and 5 µl of this dilution was used as a template for qPCR.

### **2.3.3 Stable Isotope labelling of wheat root exudates using <sup>13</sup>CO<sub>2</sub>**

Agricultural soil was sampled in July 2019, using the sampling method described in 2.2.1. Prior to use, the soil was homogenized; any organic matter, or stones larger than ~3 cm, were removed before soil was spread out to a depth of ~2 cm and dried at 21°C overnight. *T. aestivum* var. Paragon seeds were sown as described in 2.3.1,

and three additional pots remained unplanted as controls for autotrophic CO<sub>2</sub> fixation by soil microorganisms. Plants were grown in unsealed gas tight 4.25 L PVC chambers under a 12 h light / 12 h dark photoperiod at 21°C for three weeks. After three weeks, at the start of each photoperiod the chambers were purged with CO<sub>2</sub> free air (80% nitrogen, 20% oxygen, (British Oxygen Company, Guilford, UK)) and sealed before hourly pulse CO<sub>2</sub> injection. During each photoperiod three plants and three unplanted soil controls were injected with <sup>13</sup>CO<sub>2</sub> (99% Cambridge isotopes, Massachusetts, USA) and three plants were injected with <sup>12</sup>CO<sub>2</sub>. Headspace CO<sub>2</sub> was maintained at 800 ppmv (~twice atmospheric CO<sub>2</sub>). Plant CO<sub>2</sub> uptake rates were determined every four days to ensure the volume of CO<sub>2</sub> added at each one-hour interval would maintain approximately 800 ppmv. For this, headspace CO<sub>2</sub> concentrations were measured using gas chromatography every hour.

Measurements were conducted using an Agilent 7890A gas chromatography instrument, with flame ionization detector, a Poropak Q (6ft x 1/8") HP plotQ column (30 m x 0.530 mm, 40 µm film), a nickel catalyst, and a helium carrier gas. The instrument ran with the following settings: injector temperature 250°C, detector temperature 300°C, column temperature 115°C, and oven temperature 50°C. The injection volume was 100 µl and run time was five minutes (as the CO<sub>2</sub> retention time is ~3.4 mins). A standard curve was used to calculate the ppmv of CO<sub>2</sub> within samples from the peak areas. Standards of known CO<sub>2</sub> concentration were prepared in nitrogen flushed 120 ml serum vials. The volume of CO<sub>2</sub> required for injection at each one-hour interval in order to maintain 800 ppmv CO<sub>2</sub> was calculated as follows:

$$Vol\ CO_2\ (ml) = (800\ (ppmv) - \text{headspace}\ CO_2\ \text{after}\ 1\ \text{hour}\ (ppmv)) / 1000 * 4.25(L) .$$

At the end of each photoperiod, the lid for the PVC chambers were removed to prevent a build-up of CO<sub>2</sub> during the dark period. At the start of the next 12-hour photoperiod, tubes were flushed with CO<sub>2</sub> free air and headspace CO<sub>2</sub> was maintained at 800 ppmv as described. After 14 days of labelling, for all plants the bulk soil, rhizosphere, and endosphere compartments were sampled as described in 2.3.4.1. The 14-day labelling period was chosen to minimise the possibility for cross-feeding <sup>183,189,201</sup>, whilst providing sufficient time for root exudates to become isotopically labelled, and for root exudate utilising microbiota to become isotopically labelled. For all buffer recipes see Table 2.2.

## **2.3.4 Wheat root sampling and experimental methods**

### **2.3.4.1 Extraction of nucleic acids from wheat tissue and soil**

Microbial communities were analysed in the bulk soil, rhizosphere, and endosphere compartment for all plants. All three compartments were analysed from triplicate plants for each condition described. After de-potting, potted soil associated with each plant was homogenised and a bulk soil sample was taken. For all plants the phyllosphere was removed using a sterile scalpel and discarded. To analyse the rhizosphere and endosphere samples loose soil was first shaken off of the roots, then roots were washed in phosphate buffered saline (PBS). Pelleted material from this initial wash was analysed as the rhizosphere sample. To obtain the endosphere samples, remaining soil particles were washed off of the roots with PBS buffer. Then roots were soaked for 30 seconds in 70% ethanol (v/v), 5 minutes in 3% sodium hypochlorite (v/v) and washed 10 times with sterile water for surface sterilisation. To remove the rhizoplane roots were then sonicated for 20 minutes in a sonicating water bath<sup>52</sup>. After processing, all root, rhizosphere, and soil samples were snap frozen in liquid nitrogen and stored at -80°C. The frozen root material was ground up in liquid nitrogen with a pestle and mortar to a fine powder. To sample seeds five wheat seeds (each weighing ~70 mg) per replicate were surface sterilised (2.3.1) before being ground into a fine powder in liquid nitrogen with a pestle and mortar. As with root samples, material was then transferred to a lysing matrix E tube, snap frozen in liquid nitrogen and stored at -80°C. For all samples DNA was extracted using the FastDNA™ SPIN Kit for Soil (MP Biomedical) according to manufacturer's protocol with minor modifications: incubation in DNA matrix buffer was performed for 12 minutes and elution was carried out using 75 µl DNase/Pyrogen-Free Water. All DNA samples were stored at -20°C. DNA quality and yields were then assessed using a nanodrop and Qubit fluorimeter. For buffer recipes see Table 2.2.

### **2.3.4.2 Cultivation of endophytic microbes from wheat roots**

After cultivation for four weeks in either freshly sampled agricultural soil or Levingtons F2 compost, as described in 2.3.1, plants were de-potted and the phyllosphere was removed with a sterile scalpel. Then, after all soil particles has been removed by washing roots in PBS, roots were soaked for 30 seconds in 70% ethanol (v/v), 5 minutes in 3% sodium hypochlorite (v/v) and washed 10 times with sterile water for surface sterilisation. For each plant the entire root was then crushed in 2 ml 10% glycerol (v/v) with a pestle and mortar. When root material was fully

homogenised, a serial dilution was performed and 100 µl of the 10<sup>-1</sup> and 10<sup>-2</sup> dilution was plated on either R2A, R2A-O, BAZ, BAZ-O, and MAG, all supplemented with the antifungals cycloheximide (100 µg / ml) and nystatin (10 µg / ml). In a second experiment, performed in the same way, R2A, BCSA, and NA were used with vancomycin (25 µg / ml) added to inhibit the growth of gram-positive bacteria. BCSA also contained gentamycin (10 µg / ml) and polymyxin B (600 Units / ml). For media or buffer recipes see Tables 2.1 and 2.2 respectively. Plates were incubated at 30°C for ~24 h before a maximum of 100 colonies were patch plated from each medium for identification 16S rRNA gene colony PCR and sequencing.

### **2.3.5 Archaeal plant growth promotion experiments**

#### **2.3.5.1 Wheat plant growth promotion experiment**

To assess the effect of ammonia oxidising archaeon (AOA) *Nitrosocosmicus franklandus* C13 on the growth of wheat, first three 1.6 litre cultures of the strain were cultivated in *Nitrosocosmicus* medium (Table 2.1). These cultures were inoculated with *N. franklandus* C13 using an established (two-week-old) culture. To confirm that the strain was growing, *N. franklandus* C13 activity was measured via a nitrite assay, which is used as a proxy for ammonia oxidation. Over two weeks 1 ml was sampled from each culture every four days and the colorimetric Griess reaction was used as previously described in Miranda *et al.*<sup>265</sup> to measure nitrite concentrations. This confirmed that ammonia oxidation was occurring within all three cultures, which were thus concluded to be actively growing. After incubation at 37°C for 2 weeks archaeal biomass was harvested from the three 1.6 litre cultures by passing the culture medium through a 0.2 µm filter using a vacuum pump. Then the filter was then inverted, and biomass was eluted in 20 ml of MSK medium, and cells were pelleted via centrifugation. Cells were then resuspended in 1 ml MSK.

To estimate the number of archaeal cells per ml, 10 µl was taken and a 10<sup>-3</sup> dilution was performed into a final volume of 1 ml of sterile dH<sub>2</sub>O. Two drops NucBlue™ DAPI reagent were added (Thermo Fisher, MA USA) to produce a blue fluorescence where DNA was present under UV light, making the cells easier to count. The solution was mixed thoroughly before incubation for 15 minutes, then mixed again before 1 ml of the solution was loaded on to a nitrocellulose disk using a vacuum pump. Using a fluorescence microscope (excitation emission 360 nm / 460 nm), the number of cells were counted within five circular 41526.5 µm<sup>2</sup> areas. The average of these five areas was used to calculate the approximate number of archaeal cells over the whole disk, and therefore within the 1 ml 10<sup>-3</sup> dilution of the original cell-

suspension. The following equation was used, where 4839.32 is the number of times the area of the disk is divisible by the area counted:

$$(\text{mean cells per } 41526.5 \mu\text{m}^2 * 4839.32) * 10,000 = \text{cells/ml}$$

Then, the concentration of cells in the original stock was calculated and this cell suspension was diluted to a concentration of  $10^8$  cells per ml in MSK medium. Surface sterilised wheat seeds (see section 2.3.1) were then incubated for two hours in either of the control conditions (sterile water or sterile MSK medium) or in the  $10^8$  cell per ml suspension. After this, *T. aestivum* paragon var. seeds were sown in pots of pre-wetted agricultural soil, and a further 10 ml of each treatment was added into the soil surrounding the seed. After cultivation for four weeks, roots and shoots were sampled separately, and all soil was cleaned off of via 3 washes in sterile dH<sub>2</sub>O. For each experimental replicate 12 plants were sampled per treatment. The shoots and roots were then placed into separate falcon tubes, each pre-weighed with the lid, and these tubes were opened and incubated at 75°C for 10 hours to dry the tissue before they were weighed. Root or shoot dry weights were calculated by subtracting the tube weight from the total for each sample, and total dry weight was calculated as the sum of the root and shoot dry weight for each individual plant. This process was repeated in experimental triplicate, in total 36 replicates were sampled per treatment across three experimental replicates. Statistical comparisons were performed using R <sup>266</sup>, the packages dplyr <sup>267</sup> (version 1.0.4) and Dunns.test <sup>268</sup> (version 1.3.5) were used to run a Kruskal Wallis and a Dunns Test.

### **2.3.5.2 Preliminary *Arabidopsis* plant growth promotion experiment**

An initial experiment aimed to assess the impact of ammonia oxidising archaeon (AOA) *Nitrosocosmicus franklandus* C13 on the growth of *Arabidopsis thaliana* Col-0. A mature two-week old *N. franklandus* C13 culture was used to inoculate a 1.6 litres of AOA medium, and growth was monitored as described in 2.3.5.1. After two weeks of incubation at 37°C *N. franklandus* C13 cells were harvested with a vacuum pump and cell concentration was measured and normalised to  $10^8$  cells / ml as described in 2.3.5.1. *A. thaliana* Col-0 seeds were surface sterilised, for this the seeds were incubated in 70% ethanol for 2 minutes, then 20% bleach for 2 minutes before washing 5 times with sterile dH<sub>2</sub>O. After surface sterilisation seeds were incubated on an inverter for 1 hour and 30 minutes in one of three treatments, a  $10^8$  suspension of *N. franklandus* C13 in AOA medium, an uninoculated AOA medium control, or a dH<sub>2</sub>O control. After this, seeds were sown into individual pots of Levingtons F2 compost and 20 ml of the relevant treatment was added into the soil.

Plants were cultivated at 21°C under a 12 h light / 12 h dark photoperiod for 76 days before shoots were sampled and fresh weights were weighed, 12 plants were sampled per treatment. Statistical analysis was performed as described in 2.3.5.1.

## **2.4 Nucleic acid methods**

### **2.4.1 Electrophoresis**

#### **2.4.1.1 Agarose gel electrophoresis**

Agarose gel electrophoresis was performed using 1% agarose gels made with TBE buffer and 2 µg / ml ethidium bromide. DNA samples were loaded in bromophenol blue loading buffer. All samples ran alongside 1 kb plus DNA ladder (Invitrogen) and imaged using a Bio-Rad gel doc XR Imager system. For buffer recipes see Table 2.2.

#### **2.4.1.2 Denaturing Gradient Gel Electrophoresis (DGGE)**

DGGE was performed separately on the bacterial and archaeal 16S rRNA genes to screen fractions from stable isotope probing (SIP) density gradient ultracentrifugation for a change in the community in the heavy compared to the light fractions, and between the <sup>13</sup>CO<sub>2</sub> labelled heavy fractions and those of the <sup>12</sup>CO<sub>2</sub> control plants. A nested PCR approach was taken to amplify the archaeal 16S rRNA gene, the first round used primers A109F/A1000R and the second introduced a 5' GC clamp using A771F-GC/A975R. The same method was used to screen for a shift in the archaeal community across root compartments. One round of PCR was used for bacterial DGGE using the primers PRK341F-GC/518R to introduce a 5' GC clamp, and for archaeal *amoA* DGGE using CrenamoA23f/A616r. Primer sequences are available in Table 2.4. PCR conditions are indicated in Table 2.5. An 8% polyacrylamide gel was made with a denaturing gradient of 40-80% (2.8 M urea / 16% (vol/vol) formamide, to 5.6 M urea / 32% (vol/vol) formamide), with a 6% acrylamide stacking gel with 0% denaturant. 2-8 µl of PCR product was loaded per well for each sample and the gel was loaded into an electrophoresis tank filled with 1x TAE buffer. Electrophoresis was ran at 0.2 amps, 75 volts and at 60°C for 16 hours. After washing, gels were stained in the dark using 4 µl of SYBR gold nucleic acid gel stain (Invitrogen) in 400 ml 1x TAE buffer. After one hour, gels were washed twice before imaging using a Bio-Rad Gel Doc XR imager. For buffer recipes see Table 2.2.

## **2.4.2 Amplification and sequencing of 16S rRNA gene and ITS2 sequences from root and soil DNA extracts**

### **2.4.2.1 Illumina sequencing of bacterial or archaeal V3-V4 or fungal ITS2 region**

Bacterial or archaeal 16S rRNA gene sequences, or fungal ITS2 gene sequences were amplified from root, seed, or soil DNA extracts (see 2.3.4.1), see Table 2.4 for primer sequences. All 16S rRNA genes were amplified using primers specific to the archaeal (A0109F/A1000R) or bacterial (PRK341F/MPRK806R) gene. The fungal 18S ITS2 region was amplified using primers specifically targeting fungi (fITS7Fw/ITS4Rev\_2) to avoid *Triticum aestivum* ITS2 amplification. The specificity of the fungal primers was validated via shotgun cloning and sequencing (data not included). No fungal ITS2 amplicon could be obtained from the endosphere of Levington F2 compost plants. PCR conditions are indicated in Table 2.5. Before sequencing, PCR products were purified via gel extraction (see section 2.4.5.1). Samples were then sent for paired-end sequencing using an Illumina MiSeq platform at Mr DNA (Molecular Research LP, Shallowater, Texas, USA). The bacterial 16S rRNA gene was sequenced using the PRK341F/MPRK806R primers (465bp). The archaeal 16S rRNA gene was sequenced using the A0349F/A0519R primers (170bp). The fungal ITS2 region was sequenced with the fITS7Fw/ITS4Rev\_2 primers (350bp). For metabarcoding experiments samples were submitted to Mr DNA in five separate runs: (1) all samples for field grown stem elongation growth phase wheat and the soil type experiment (sections 3.1, 4.1, and 4.2), (2) seed endosphere fungal community (section 4.4), (3) Senescent wheat (section 3.2), (4) wheat varieties experiment (section 4.3), (5) seed endosphere bacterial community (section 4.4). For runs 1-3 a sequencing depth of 20K was used. For run four a sequencing depth of 40K was used, and for endosphere samples peptide nucleic acid (PNA) blockers against chloroplast and mitochondria sequences were used to minimise amplification of host derived sequences. For run five both PNA blockers were used, and sequencing was performed to a depth of 120K. For stable isotope probing experiments sequencing was performed in two runs; (1) All SIP pooled fraction samples (see Table 2.6), sequenced to a depth of 50K and with PNA blockers for endosphere samples (section 5.3), (2) repeat of SIP pooled fraction endosphere samples, and unfractionated samples from the SIP experiment (section 5.3). All SIP experiment sequencing was performed to a depth of 50K and using PNA blockers for endosphere samples. See section 2.6.1 for data analysis methodology.

#### **2.4.2.2 PacBio SMRT closed circular sequencing (CSS) of bacterial V1-V9 region**

Full length bacterial 16S rRNA gene sequencing was performed on amplicons from paragon var. wheat seeds, and from the endosphere, rhizosphere, and bulk soil compartments using universal bacterial primers 27f/1492r (Table 2.4). Sequencing was performed at Mr DNA (Molecular Research LP, Shallowater, Texas, USA) using a PacBio Sequel SMRT CSS platform. For seed and endosphere samples PNA blockers were used, and sequencing was performed to a depth of 20K. For bulk soil and rhizosphere samples sequencing was performed to a depth of 10K. See section 2.6.2 for data analysis methodology.

#### **2.4.2.3 Amplification and sequencing of full length 16S rRNA or 18S rRNA gene sequences for isolate identification or the generation of qPCR standards**

Full length bacterial 16S rRNA gene sequences were amplified using generic primers (27F/1492R), whereas for archaea, specific primers were used (A0109F/A1000R). Full length fungal 18S rRNA gene sequences were amplified using primers SSUaF/1510R. For qPCR standards, amplification was performed using a soil DNA extract for archaeal and bacterial standards. For the fungal assay colony PCR was performed on *Saccharomyces cerevisiae*; after cultivation overnight in LB medium, at 37°C with shaking (250 rpm), the culture was diluted 1:10 (vol/vol) in sterile dH<sub>2</sub>O and boiled at 100°C for 20 minutes, then 5 µl of this was used as a PCR template in 50 µl reaction. After purification via gel extraction (see section 2.4.5.1) these sequences were then cloned into the Promega pGEM®-T Easy Vector system using the manufacturers protocol. The correct sequence was validated by Sanger sequencing (Eurofins Genomics, Germany) using the same primers. For isolate 16S sequencing, bacterial 16S rRNA gene sequences were amplified in the same way, after cultivation overnight in LB medium at 30°C with shaking (250 rpm). For all PCR conditions see Table 2.5 and for primer sequences see Table 2.4.

#### **2.4.3 End-point PCR screening for *Burkholderiaceae* family endophytic bacteria**

Primers touted to be *Burkholderia* specific (Burk3f/BurkR)<sup>269</sup> were used to initially screen root endosphere isolates for strains belonging to the *Burkholderia* genus. This experiment was performed prior to the long-read amplicon sequencing experiment which revealed that *Burkholderia* were not the *Burkholderiaceae* family

genus present within the roots. After cultivation on LB medium, one colony for each isolate was picked and suspended in 100 µl sterile dH<sub>2</sub>O and boiled at 100°C for 30 minutes, afterwards 1 µl was used as template in a 10 µl PCR reaction. As a control, generic 16S rRNA gene primers (27F/1492R) were used to verify the template quality. Prior to the screen, the primers had previously been tested using *Streptomyces venezuelae* or *Streptomyces coelicolor* genomic DNA, and *Escherichia coli* and methicillin resistant *Staphylococcus aureus* colony PCRs (sampled in the same way) to validate specificity. For all PCR conditions see Table 2.5 and for primer sequences see Table 2.4.

#### 2.4.4 Real time quantitative PCR

The abundance of bacterial or archaeal 16S rRNA genes and of fungal 18S rRNA genes within samples was determined by qPCR amplification of these genes from DNA extracts. Bacterial 16S rRNA abundance was quantified using bacteria-specific primers Com1F/769r, as previously described<sup>270</sup>. Archaeal 16S rRNA gene abundance was quantified using the archaeal specific A771f/A957r primers, as previously described<sup>232</sup>. Fungi-specific primers FR1F/FF390R, as previously described<sup>271</sup>, were used to quantify 18S rRNA gene abundance and examine <sup>13</sup>C labelling of the fungal community for the SIP fractions. Primer sequences are presented in Table 2.4. The qPCR was performed using the Applied Biosystems QuantStudio 1 Real-Time PCR System (Applied Biosystems, Warrington, UK) with the New England Biolabs SYBR Green Luna® Universal qPCR Master Mix (New England Biolabs, Hitchin, UK). PCR mixtures and cycling conditions are described in Table 2.5. Bacterial, fungal, and archaeal qPCR standard vectors were prepared as described (2.4.5.2), to prepare qPCR standards, amplification of the qPCR standard was performed off of these vectors and double purified using PCR clean-up (2.4.6.2). Then, the standard was diluted from 2x10<sup>7</sup> to 2x10<sup>0</sup> copies / µl in duplicate and ran alongside all qPCR assays. Ct values from standard dilutions were plotted as a standard curve and used to calculate 16S/18S rRNA gene copies / 50 ng DNA extract. Amplification efficiencies ranged from 90.9% to 107% with R<sup>2</sup> > 0.98 for all standard curve regressions. All test samples were normalised to 50 ng of template DNA per reaction and ran in biological triplicate. PCR products were all analysed by both melt curves and agarose gel electrophoresis which confirmed amplification of only one product of the expected size. For statistical comparison of the average 16S rRNA or 18S rRNA gene copy number between samples ANOVA and linear models, followed by Tukey post-hoc was run in R<sup>266</sup>. For melt curves see Supplementary Figures S.14-S.16.

**Table 2.4. Primer sequences**

Primer	Target	Uses	Sequence	Reference
PRK341F-GC	Bacterial 16S	DGGE PCR Amplification	CGCCCGCCGCGCGCGGGCGGG CGGGGCGGGGGCACGGGGGG CCTACGGGAGGCAGCAG	(Muyzer <i>et al.</i> , 1993) <sup>272</sup>
518R	Bacterial 16S	DGGE PCR Amplification	ATTACCGCGGCTGCTGG	(Muyzer <i>et al.</i> , 1993) <sup>272</sup>
A771F-GC	Archaeal 16S	DGGE PCR Amplification	CGCCCGCCGCGCGCGGGCGGG CGGGGCGGGGGCACGGGGGG ACGGTGAGGGATGAAAGCT	(Ochsenreiter <i>et al.</i> , 2003) <sup>273</sup>
A957R	Archaeal 16S	DGGE PCR Amplification, qPCR	CGGCGTTGACTCCAATTG	(Ochsenreiter <i>et al.</i> , 2003) <sup>273</sup>
A109F	Archaeal 16S	PCR amplification, qPCR standard amplification	ACKGCTCAGTAACACGT	(Großkopf <i>et al.</i> , 1998) <sup>277</sup>
A1000R	Archaeal 16S	PCR amplification, qPCR standard amplification	GGCCATGCACYWCYTCTC	(Gantner <i>et al.</i> , 2011) <sup>274</sup>
PRK341F	Bacterial 16S	PCR amplification/ Sequencing, qPCR standard amplification	CCTACGGGRBGCASCAG	(Yu <i>et al.</i> , 2005) <sup>275</sup>
MPRK806R	Bacterial 16S	PCR amplification/ Sequencing, qPCR standard amplification	GGACTACNNGGTATCTAAT	(Yu <i>et al.</i> , 2005) <sup>275</sup>
fITS7F	Fungal ITS2	PCR Amplification/ sequencing	GTGARTCATCGAATCTTTG	(Ihrmark <i>et al.</i> , 2012) <sup>276</sup>
ITS4R_2	Fungal ITS2	PCR Amplification/ sequencing	TCCTCCGCTTATTGATATGC	(White <i>et al.</i> , 1990) <sup>277</sup>
A771F	Archaeal 16S	qPCR	ACGGTGAGGGATGAAAGCT	(Ochsenreiter <i>et al.</i> , 2003) <sup>273</sup>
FR1Fw	Fungal 18S	qPCR	AICCATTCAATCGGTAIT	(Vainio <i>et al.</i> , 2000) <sup>278</sup>
FF390Rev	Fungal 18S	qPCR	CGATAACGAACGAGACCT	(Vainio <i>et al.</i> , 2000) <sup>278</sup>
Com1F	Bacterial 16S	qPCR	CAGCAGCCGCGGTAATAC	(Fredriksson <i>et al.</i> , 2013) <sup>279</sup>
769R	Bacterial 16S	qPCR	ATCCTGTTTGMTMCCCVCR	(Rastogi <i>et al.</i> , 2010) <sup>280</sup>
F18SS03-F	Fungal 18S	qPCR standard amplification	AGATCCTGAGGCCTCACTA	This Study
F18SS03-R	Fungal 18S	qPCR standard amplification	GCCGTTCTTAGTTGGTGGAG	This Study
18SF	Fungal 18S	Full length 18S rRNA gene cloning and sequencing	GTA A A A G T C G T A A C A A G G T T T C	(Findley <i>et al.</i> , 2013) <sup>281</sup>

1510R	Fungal 18S	Full length 18S rRNA gene cloning and sequencing	CCTTCYGCAGGTTACCTAC	Amaral-Zettler <i>et al.</i> , 2009) <sup>282</sup>
18S1A	Generic 18S	Generic 18S rRNA gene amplification	AACCTGGTTGATCCTGCCAGT	(Wang <i>et al.</i> , 2014) <sup>283</sup>
18S564R	Generic 18S	Generic 18S rRNA gene amplification	GGCACCAGACTTGCCCTC	(Wang <i>et al.</i> , 2014) <sup>283</sup>
A0349F	Archaeal 16S	Sequencing	GYGCASCAGKCGMGAAW	(Takai <i>et al.</i> , 2000) <sup>284</sup>
A0519R	Archaeal 16S	Sequencing	TTACCGCGGCKGCTG	(Takai <i>et al.</i> , 2000) <sup>284</sup>
CrenamoA2 3f	Archaeal <i>amoA</i>	DGGE	ATGGTCTGGCTWAGACG	(Tourna <i>et al.</i> , 2008) <sup>285</sup>
CrenamoA6 16r	Archaeal <i>amoA</i>	DGGE	GCCATCCATCTGTATGTCCA	(Tourna <i>et al.</i> , 2008) <sup>285</sup>
27f	Bacterial 16S	Full length 16S rRNA gene sequencing	AGAGTTTGATCMTGGCTCAG	(Heuer <i>et al.</i> , 1997) <sup>286</sup>
1492r	Bacterial 16S	Full length 16S rRNA gene cloning and sequencing	TACGGYTACCTTGTAGGACTT	(Heuer <i>et al.</i> , 1997) <sup>286</sup>
Burk3f	<i>Burkhold eria</i> 16S	Isolate PCR screen	CTGCGAAAGCCGGAT	(Salles <i>et al.</i> , 2001) <sup>269</sup>
Burkr	<i>Burkhold eria</i> 16S	Isolate PCR screen	TGCCATACTCTAGCYGCG	(Salles <i>et al.</i> , 2001) <sup>269</sup>

**Table 2.5. PCR Conditions for metabarcoding, qPCR & DGGE PCRs**

PCR Component	Volume (µl)
2x PCRBio BioMix™ red, containing BIOTAQ™ DNA Polymerase, or 2x PCRBio Ultra mix, containing Ultra DNA Polymerase	12.5
Forward or reverse primer (10mM stock)	1.25
Template DNA	2.5 (DNA extract) 1.25 (Round 1 product for nested PCR)
Sterile dH <sub>2</sub> O	Up to 25
<b>qPCR mix</b>	
2x SYBR Green Luna® Universal qPCR Master Mix	10
Forward or reverse primer (10mM stock)	0.5
DNA template (20ng / µl stock)	5
Sterile MilliQ dH <sub>2</sub> O	Up to 20
<b>Thermocycler programs</b>	
A0109F/A1000R Archaeal 16S rRNA gene For sequencing & round one of DGGE	<ol style="list-style-type: none"> <li>1. 95°C for 1 minute</li> <li>2. 35x cycles of 95°C for 30 seconds, 59°C for 30 seconds, 72°C for 45 seconds</li> <li>3. 72°C for 1 minute</li> </ol>
A771F-GC/A957R Archaeal 16S rRNA gene For round two of DGGE	<ol style="list-style-type: none"> <li>1. 95°C for 1 minute</li> <li>2. 35 cycles of 94°C for 30 seconds, 55°C for 30 seconds, 72°C for 1 minute</li> <li>3. 72°C for 10 minutes</li> </ol>
PRK341F/MPRK806R or fITS7F/ITS4R Bacterial 16S rRNA gene or Fungal ITS2 region For sequencing	<ol style="list-style-type: none"> <li>1. 95°C for 1 minute</li> <li>2. 30 cycles of 95°C for 15 seconds, 55°C for 15 seconds, 72°C for 15 seconds</li> <li>3. 72°C for 10 minutes</li> </ol>
27f/1492r or 18Sf/151r For full length bacterial 16S rRNA or fungal 18S rRNA gene amplification for cloning or sequencing	<ol style="list-style-type: none"> <li>1. 95°C for 1 minute</li> <li>2. 30 cycles of 95°C for 15 seconds, 55°C for 30 seconds, 72°C for 15 seconds</li> <li>3. 72°C for 10 minutes</li> </ol>
A771F/A957R, FR1Fw/FF390Rev or Com1F/769R qPCR assays	<ol style="list-style-type: none"> <li>1. 95°C for 10 minutes</li> <li>2. 35 cycles of 95°C for 15 seconds and 60°C annealing/ extension/ read step for 30 seconds</li> </ol>
Burk3f/Burkr For isolate <i>Burkholderia</i> screen Touchdown PCR	<ol style="list-style-type: none"> <li>1. 94°C for 4 minutes</li> <li>2. 5 cycles of 94°C for 60 seconds, 62°C for 90 seconds, 72°C for 30 seconds</li> <li>3. 5 cycles of 94°C for 60 seconds, 61°C for 90 seconds, 72°C for 30 seconds</li> <li>4. 25 cycles of 94°C for 60 seconds, 58°C for 90 seconds, 72°C for 30 seconds</li> <li>5. 72°C for 10 minutes</li> </ol>

## **2.4.5 Purification of PCR products and plasmids**

### **2.4.5.1 Gel purification**

After the correct size was verified using the ladder, PCR products were excised from the gel using a clean scalpel and placed in a pre-weighed 1.5 ml microcentrifuge tube. Then, all gel purification was performed using the manufacturers protocol for the Qiagen QIAquick gel extraction kit and DNA was eluted using 20-50  $\mu$ l sterile dH<sub>2</sub>O pre-warmed to 55°C. Quantity and quality was assessed using a Nanodrop spectrophotometer (Thermo Fisher) and a Qubit fluorometer dsDNA high sensitivity assay (Invitrogen).

### **2.4.5.2 PCR purification**

Where fragment size selection was not needed, a Qiagen QIAquick PCR purification kit was used according to the manufacturers protocol. DNA was eluted using 50-100  $\mu$ l sterile dH<sub>2</sub>O pre-warmed to 55°C. Quantity and quality was assessed using a Nanodrop spectrophotometer and a Qubit fluorometer dsDNA high sensitivity assay, or a dsDNA broad range assay (Invitrogen).

### **2.4.5.3 Plasmid purification**

*E. coli* TOP10 transformants were cultivated overnight at 37°C with shaking (250 rpm) in LB medium with the appropriate antibiotic to maintain selection for the plasmid. Then, 1 ml of culture was taken for plasmid preparation using the Promega Wizard Plus SV Miniprep kit according to the manufacturers protocol. DNA was eluted using 50-100  $\mu$ l sterile dH<sub>2</sub>O pre-warmed to 55°C. Quantity and quality was assessed using a Nanodrop spectrophotometer and a Qubit fluorometer dsDNA high sensitivity assay, or a dsDNA broad range assay (Invitrogen).

## **2.4.6 Density gradient ultracentrifugation**

Density gradient ultracentrifugation was used to separate <sup>13</sup>C labelled DNA from <sup>12</sup>C DNA as previously described by Neufeld and colleagues <sup>189</sup>. For each sample 700 ng of DNA was mixed with a 7.163 M CsCl solution and gradient buffer to a final measured buoyant density of 1.725 g / ml<sup>-1</sup>. Buoyant density was determined via the refractive index using a refractometer (Reichert Analytical Instruments, NY, USA). Samples were loaded into polyallomer quick seal centrifuge tubes (Beckman Coulter) and heat-sealed. Tubes were placed into a Vti 65.2 rotor (Beckman-Coulter) and centrifuged for 62 hours at 44,100 rpm (~177,000 gav) and 20°C under a vacuum. Samples were then fractionated by piercing the bottom and the top of the

ultracentrifuge tube with a 0.6 mm sterile needle. Through the top needle dH<sub>2</sub>O was pumped into the centrifuge tube at a rate of 450 µl per minute. As the gradient was displaced it was able to drip into 1.5ml microcentrifuge tubes. Fractions were collected until the water had fully displaced the gradient solution; this resulted in twelve ~450 µl fractions. The DNA was precipitated from fractions by adding 4 µl of Co-precipitant Pink Linear Polyacrylamide (Bioline) and 2 volumes of PEG-NaCl solution to each fraction, followed by an overnight incubation at 4°C. Fractions were then centrifuged at 21,130 g for 30 minutes and the supernatant was discarded. The DNA pellet was then washed in 500 µl 70% EtOH and centrifuged at 21,130 g for 10 minutes. The resulting pellet was air-dried and resuspended in 30 µl sterile dH<sub>2</sub>O. Fractions were then stored at -20°C and the extent of <sup>13</sup>C labelling for the bacterial, archaeal, and fungal community was analysed as described using DGGE and qPCR. Fractions were then pooled prior to sequencing (see Table 2.6 for details), and sequencing was performed as described in 2.4.2.1. For buffer recipes see Table 2.2.

**Table 2.6. SIP fractions used for sequencing**

<b>Origin of fractions</b>	<b>Fraction classification</b>	<b>Pooled fractions</b>	<b>fraction density range (g / ml<sup>-1</sup>)</b>
Endosphere A	12H	7,8,9	1.7086-1.7204
	12L	11,12	1.3956-1.6910
	13H	7,8,9	1.7098-1.7227
	13L	11,12	1.6004-1.6992
Endosphere B	12H	7,8,9	1.7098-1.7204
	12L	11,12	1.5097-1.6957
	13H	7,8,9	1.7098-1.7216
	13L	11,12	1.5745-1.6992
Endosphere C	12H	7,8,9	1.7086-1.7204
	12L	11,12	1.5427-1.6957
	13H	7,8,9	1.7110-1.7216
	13L	11,12	1.6310-1.7004
Rhizosphere A	12H	7,8	1.7169-1.7227
	12L	10,11	1.6992-1.7051
	13H	7,8	1.7169-1.7216
	13L	10,11	1.6992-1.7039
Rhizosphere B	12H	7,8	1.7157-1.7204
	12L	10,11	1.7004-1.7051
	13H	7,8	1.7157-1.7216
	13L	10,11	1.7004-1.7051
Rhizosphere C	12H	7,8	1.7169-1.7227
	12L	10,11	1.7016-1.7051
	13H	7,8	1.7157-1.7216
	13L	10,11	1.7004-1.7051
Unplanted A	13H	8,9	1.7086-1.7145
	13L	11	1.6992
Unplanted B	13H	8	1.7145
	13L	10,11	1.7004-1.7039
Unplanted C	13H	8,9	1.7098-1.7157
	13L	11,12	1.6745-1.7004

## 2.5 Microbiological techniques

### 2.5.1 Endophyte spot bioactivity screen

To screen endophytes for bioactivity in vitro first strains of interest were cultivated individually in liquid LB medium overnight at 30°C with shaking (250 rpm). Three indicator strains, *Candida albicans* (fungal), *Bacillus subtilis* (Gram-positive bacteria), and *Enterobacter aerogenes* (Gram-negative bacteria), were cultivated overnight in LB medium overnight at 37°C with shaking (250 rpm). After this a sterile cotton bud was soaked with culture for each indicator strain and spread over a separate LB agar plate for each indicator. Once dry, 5 µl of culture from each strain of interest was spotted on to the LB agar spread with each of the three indicator strains, leaving an approximately 1 cm gap between each strain. In addition, a positive control was also spotted on to the plates in the form of an antibiotic, cycloheximide (10 mg / ml) for *C. albicans*, and hygromycin (50 mg / ml) for *E. aerogenes* and *B. subtilis*. Plates were then incubated overnight at 30°C. Strains which showed activity had a zone of clearing around the colony, and were re-tested individually in biological triplicate, in the same way as described for the relevant indicator strains.

### 2.5.2 Test for root endophytes interactions with *Streptomyces*

Onto either NA medium or MM, 3 µl of spores from *Streptomyces* endophyte strain CRS3 were spotted around 2 cm from the edge of the plate. After the spores were dry, plates were incubated at 30°C. After three days all *Burkholderiaceae* and *Pseudomonas* isolates were cultivated in LB medium at 30°C overnight with shaking at 250 rpm, except for RR205 and RR307 which were cultivated at 37°C with shaking at 250 rpm. Then, after CRS3 had been incubated at 30°C for three days and all *Burkholderiaceae* and *Pseudomonas* isolates had been cultivated in liquid medium overnight, 3 µl of culture medium was spotted onto the plate opposite the *Streptomyces* CRS3 colony, approximately 2 cm from the edge of the plate, with three replicates per isolate. Additionally, each *Burkholderiaceae* and *Pseudomonas* isolate was inoculated on NA without *Streptomyces*. These plates were then incubated at 30°C for four to 11 days before imaging.

### 2.5.3 Wheat take-all fungus bioassays

Take-all fungus (*Gaeumannomyces tritici*<sup>287</sup>) was cultivated on PDA for 5 days before an actively growing plug was transferred to a new plate of PDA.

*Pseudomonas* genus root isolates were cultivated in LB medium at 30°C overnight

with shaking at 250rpm. After the take-all fungus has been cultivated for three days at 21°C, 3 µl of culture was spotted on to triplicate plates approximately 2.5 cm away from the take-all plug, and three take-all plates remained uninoculated as a control. After incubation for three days at 21°C, plates were imaged.

#### **2.5.4 Ammonia oxidising archaea enrichment culture**

To test the ability of AOA *Nitrosocosmicus franklandus* C13 to survive in MSK plant growth medium, four 800 ml cultures were inoculated with 5 ml of an established two-week old culture. Each 800 ml culture had slightly varied conditions, 5V was half strength MSK supplemented with 0.8 ml AOA medium vitamins, 5O was half strength MSK without any vitamin supplement, FV was MSK medium supplemented 0.8 ml AOA medium vitamins, and FO was MSK medium. The pH of MSK medium was 6.31, the pH of half strength MSK medium was 6.34. For media recipes see Table 2.1 and for vitamin solution recipe see Table 2.3. For half strength medium (FV or FO) the 10x salts solution was diluted 1:1 in sterile dH<sub>2</sub>O prior to autoclaving, and half the concentration of all additives was added. After inoculation the cultures were incubated at 37°C for 40 days, and nitrite concentration was measured to assess ammonia oxidising activity, which was used as a proxy for growth. This was performed as described in 2.3.5.1, a zero measurement was taken prior to incubation then further nitrite measurements were made after five, 14, 19, 25 and 40 days.

## **2.6 Computational and statistical analysis**

### **2.6.1 Illumina amplicon sequencing analysis**

As detailed in section 2.4.2.1, all 16S rRNA gene and ITS2 region sequencing was performed using an Illumina MiSeq platform at Mr DNA (Molecular Research LP, Shallowater, Texas, USA). When received, all sequencing reads were further processed using the software package quantitative insights into microbial ecology 2 (Qiime2) version 2019.7<sup>288</sup>. Paired-end sequencing reads were demultiplexed and then quality filtered and denoised using the DADA2 plugin version 1.14<sup>289</sup>. Reads were trimmed to remove the first 17-20 base pairs (primer dependent, see Table 2.7) and truncated to 150-230 base pairs to remove low quality base calls (dependent on read quality and amplicon length, see Table 2.7). Chimeras were removed using the consensus method. The Dada2 denoising algorithm<sup>289</sup> was used to generate amplicon sequence variants (ASV's) for analysis, taking full advantage

of the fine-scale variation that is detectable using modern sequencing methods, which can be lost when reads are clustered into operational taxonomic units (OTUs)<sup>289</sup>. Default settings were used for all other analyses. For taxonomic assignments Bayesian bacterial and archaeal 16S sequence classifiers were trained against the SILVA<sup>290</sup> database version 128 using a 97% similarity cut off. For the fungal ITS2 reads, the bayesian sequence classifier was trained against the UNITE<sup>291</sup> database version 8.0 using a 97% similarity cut-off. Taxonomy-based filtering was performed to remove contaminating mitochondrial, chloroplast and *Triticum* (Supplementary Table S.1), remaining sequences were used for all further analyses. Taxonomy-based filtering was not required for the fungal dataset. For all datasets, taxonomic identification was validated via an National Centre for Biotechnology Information (NCBI) basic local alignment search tool (BLAST)<sup>292</sup> search, which verified taxonomic identification for the top three most abundant taxa.

Statistical analysis was performed using R version 3.6.2<sup>266</sup>. The package vegan version 2.5-7<sup>293</sup> was used to calculate Bray Curtis dissimilarities and conduct similarity percentages breakdown analysis (SIMPER<sup>294</sup>). Bray Curtis dissimilarity matrices were selected as they ignore taxonomic affiliation and can present community shifts based on the fine-scale variation detectable using ASV based analysis. UniFrac distances could also have been used, however as the questions posed by this work related to compositional differences, and not specifically phylogenetic differences, the Bray Curtis method was selected as the most appropriate method. Permutational Multivariate Analysis of Variance (permanova) analyses were conducted using Bray Curtis dissimilarity matrices and the adonis function in vegan. Bray Curtis dissimilarities were also used for principle coordinate analysis (PCoA) which was performed using the packages phyloseq version 1.3<sup>295</sup> and plyr. Differential abundance analysis was performed using DESeq2 in the package microbiomeSeq version 0.1<sup>296</sup>. Analysis of compositions of microbiomes with bias correction (ANCOM-BC) analysis was performed using phyloseq version 1.3<sup>295</sup> and ANCOMBC version 1.1.4<sup>297</sup>. For the variety experiment (Chapter 4.3), due to the greater sequencing depth achieved (Supplementary Table S.1) prevalence based filtering was performed prior to all statistical analysis using the package Microbiome, version 1.12<sup>298</sup>. Prevalence based filtering aimed to remove low abundance taxa likely to be sequencing artefacts; this removed taxa with fewer than an average of 50 reads per sample (the abundance threshold), and which appeared in just two or fewer samples (the prevalence threshold). For DESeq2 analysis, given the low number of reads which remained in some samples after

taxonomy-based filtering within the soil type experiment (Chapter 4.2) (see Supplementary Table S.1 for read counts), a base mean cut off of 200 was applied to the field and pot metabarcoding experiments. For the SIP experiment a base mean cut off of 400 was applied. These cut offs were used to eliminate any possible false positives resulting from low sequencing depth. For the field and pot metabarcoding experiments (including the senescent wheat dataset), if a taxon had a base mean > 200 and a significant p-value in one or more comparison, data for that taxon was retained for all comparisons and included in the figures. For details see Supplementary Tables S.2 – S.8

**Table 2.7. DADA2 settings**

<b>Amplicon</b>	<b>p-trim-left-f</b>	<b>p-trim-left-r</b>	<b>p-trunc-len-f/r</b>
A0349F/A0519R Archaeal 16S	17	15	120
PRK341F/MPRK806R Bacterial 16S	17	20	230
fITS7F/ITS4R Fungal ITS2	19	20	195

## **2.6.2 PacBio full length 16S rRNA gene sequencing analysis**

As detailed in section 2.4.2.2, full length bacterial 16S rRNA gene was performed using a PacBio sequel platform at Mr DNA (Molecular Research LP, Shallowater, Texas, USA), utilising the SMRT closed circular sequencing (CSS) protocol. When received, all sequencing reads were first processed using lima version 2.1.0 to demultiplex sequences and to trim barcode and universal PacBio adapter sequences. Then, the dada2 pipeline was used to process the reads using R packages dada2 version 1.18<sup>289</sup>, Biostrings version 2.58<sup>299</sup>, ShortRead version 1.48<sup>300</sup>, ggplot2 version 3.3.5<sup>301</sup>, reshape2 1.4.4<sup>302</sup>, gridExtra version 2.3<sup>303</sup>, and phyloseq version 1.34<sup>295</sup> for R version 4.0.3<sup>266</sup>. This pipeline was used to remove primer sequences and short sequence reads, to perform the dereplication and denoising steps using pseudo-pooling, for error and chimera identification and removal (consensus method), and to perform taxonomic assignments using the SILVA<sup>290</sup> database version 1.38.1, which incorporates the taxonomy defined by the genome taxonomy database (GTDB)<sup>304</sup>. After analysing some samples individually, it became clear that the data was insufficient for analysis to be performed on

individual samples as too few reads remained after the quality filtering or denoising step (Table 2.8), and the data failed the error check step (Supplementary Figures S.5-S.10). As this analysis was computationally demanding and time consuming to run (> 3-4 days per sample) the decision was made to pool reads from the three replicates for each sample and process each sample type as a single file in order to recover as many 16S rRNA gene sequences from the data as possible. Pooled endosphere, rhizosphere, and bulk soil samples all passed the quality check (Supplementary Figures S.11-S.12). For the rhizosphere, bulk soil, and seed samples this failed to yield a high quantity of reads, and the seed data failed the quality check (Supplementary Figure S.13). Identifiable 16S rRNA gene sequences were still recovered for the bulk soil and rhizosphere compartments. Given the highly unequal depth after taxonomic filtering (Table 2.8) comparisons between the community composition in each compartment were not made, and instead this data was used for the genus to species level identification of microbial taxa within the rhizosphere and endosphere.

For each bacterial family of interest, a Fasta file was generated containing all of the unique 16S rRNA gene sequences. In addition, the 16S rRNA gene sequence for any isolates contained within that group were also added to this file (see Chapter 6), along with example sequences for any specific species or genera identified, and for rooting taxa. All example sequences were acquired from the NCBI database. Rooting taxa were selected from a closely related family or genus within the same order or family as the groups of interest (these are described within the figure legends in chapter four, figures 4.11 to 4.16). Then, Clustalw version 2.1 was used to produce an alignment for these sequences, to produce a guide tree, and to produce percent sequence identity matrices<sup>305</sup>. For the groups *Streptomycetaceae* and *Pseudomonadaceae*, Clustalw failed to produce an identity matrix, so for these two groups a pairwise BLAST search<sup>292</sup> was ran to compare all individual 16S rRNA gene sequences and to produce a percent identity matrix. The tree data was loaded into the tree making tool FigTree version 1.4 which was used to produce the final rooted trees presented in chapter four.

**Table 2.8. Reads remaining in pooled or un-pooled PacBio CSS samples after demultiplexing using lima, and after each step in the dada2 analysis pipeline**

Sample	Demultiplexing	Primer trimming	Quality filtering	Denoising	Taxonomy based filtering
Endosphere pooled	136,808	136,611	83,057	52,292	12,635
Rhizosphere pooled	35,907	32,188	17,658	118	118
Bulk soil pooled	37,671	33,483	18,662	50	50
Seeds pooled	3,158	951	39	1	0
<b>Un-pooled samples</b>					
Endo_1	136,246	136,148	75,950	48,493	
Endo_2					
Endo_3	147	110	5		
Rhizo_1					
Rhizo_2					
Rhizo_3					
BulkSoil_1	14,701	13,270	7,609	8	
BulkSoil_2	18,550	17,313	10,671		
BulkSoil_3					
Seed_1	223	168	7		
Seed_2	1711	594	18		
Seed_3	1224	189	11		
= n.a, step was not ran for this sample *					

## Chapter 3. Cross-domain compositional analysis of the wheat root microbiome

Sections 3.1, 3.2 and 3.3.1 have been published previously <sup>198</sup>.

### Aims

The experiments reported in this chapter aimed to profile the microbial community associated with the roots of Paragon var. *Triticum aestivum*, and to assess the effect of developmental senescence on the community. While many studies have profiled the wheat root microbiome under a range of conditions <sup>55,63,72,94–99,109</sup>, few studies have profiled the endosphere community, inside the root <sup>108,109</sup>. This chapter was aimed at addressing this knowledge gap by profiling the endosphere community for wheat in addition to the rhizosphere. To our knowledge the root associated microbiome has not been studied for wheat during developmental senescence. Developmental senescence is the final stage in development, where nutrients (particularly nitrogen) are remobilised from the plant tissues to the developing grain. At this point, the plants are no longer green or actively growing, and it is likely that root exudation profiles are affected (particularly nitrogen containing root exuded compounds; plants can exude up to 15% of their available nitrogen from the roots when actively growing <sup>306</sup>). The process of developmental senescence is a key determinant of yields and grain nutrient content, thus we felt it important to understand how the microbial community is changing during this process. There is some limited evidence that archaea may form beneficial symbioses with terrestrial plants <sup>236</sup>, this group however are often overlooked in microbial community surveys. This chapter aimed to provide an inclusive overview of the taxa present within the root microbiome by assessing archaeal diversity, in addition to the fungal and the bacterial diversity.

### Results

#### 3.1 Profiling the microbiome associated with field grown wheat

To gain initial insights into the microbial communities associated with wheat roots, Paragon var. *Triticum aestivum* plants were sampled from the field during the stem elongation growth phase. The diversity of microbes in the bulk soil, rhizosphere, and endosphere compartments was investigated using 16S rRNA gene (for bacteria and archaea) or ITS2 (for fungi) metabarcoding. Using archaea-specific primers <sup>227,274</sup>

we were able to amplify archaeal 16S rRNA gene sequences from the root endosphere (Figure 3.1), demonstrating that archaea were present within the wheat root endosphere.

The bacterial and fungal communities differed significantly across compartments (bacterial permanova:  $R^2 = 0.8$ ,  $p < 0.01$ ; fungal permanova:  $R^2 = 0.63$ ,  $p < 0.01$ ). This was particularly the case for the rhizosphere and endosphere compartments when compared to bulk soil, as demonstrated by principal coordinates analysis (PCoA) (Figure 3.2, A). Compositional profiles did not indicate a strong shift in the archaeal community across compartments at the family level (Figure 3.3, C), but at the community level statistical analysis showed that there was a significant effect of compartment on the archaeal community composition (archaeal permanova:  $R^2 = 0.66$ ,  $p < 0.01$ ), and PCoA indicated that differences in the endosphere may mostly be responsible for this shift (Figure 3.2, A2). This shows that there was a shift in the archaeal community across root compartments, though the changes in the community were subtle.

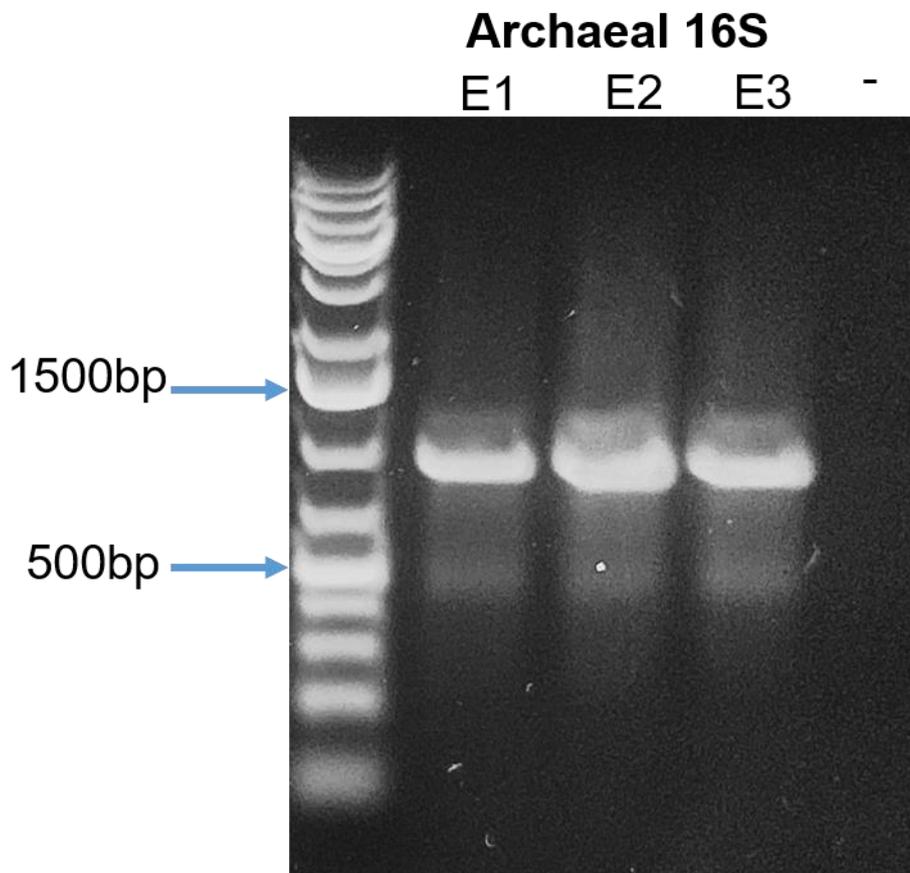


Figure 3.1. Agarose gel electrophoresis showing amplification of archaeal 16S rRNA gene sequences from the wheat root endosphere. Bands in lanes E1, E2, and E3 show archaeal 16S rRNA gene amplification from three replicate paragon var. endosphere samples, predicted size 1000bp. The lane marked "-" shows a negative control. The ladder was gene ruler 1KB+ (Invitrogen).

For the bacterial community the family *Streptomycetaceae* showed the greatest average relative abundance in the endosphere (25.12%), followed by *Burkholderiaceae* (11.99%), and *Sphingobacteriaceae* (7.75%). In the rhizosphere the relative abundance of *Streptomycetaceae* was much lower (2.58%), while *Micrococcaceae* were most abundant family (8.43%), followed by *Burkholderiaceae* (7.41%), and *Sphingobacteriaceae* (6.58%) (Figure 3.3, A). The fungal endosphere community was dominated by the Xyariales order (32.9%), followed by the class *Sordariomycetes* (14.33%), then the *Metarhizium* (10.44%). For the rhizosphere, however *Metarhizium* showed the greatest relative abundance (27.36%), followed by the Chaetothyriales order (12.32%) and the *Sordariomycetes* (9.23%) (Figure 3.3, B). The archaeal community was dominated by the AOA family *Nitrososphaeraceae* (endosphere 89.77%, rhizosphere 81.55%) (Figure 3.3, C).

Differential abundance analysis demonstrated that the abundance of fourteen bacterial families, including *Streptomycetaceae*, *Burkholderiaceae*, and *Sphingobacteriaceae*, increased significantly within the rhizosphere and/or the endosphere relative to the bulk soil (Figure 3.5 A). The families *Streptomycetaceae* (16.4% contribution,  $p < 0.01$ ) and *Burkholderiaceae* (6.1% contribution,  $p < 0.01$ ) were the two most significant contributors to the bacterial community shift as confirmed by SIMPER analysis (see Table 3.1 for full outputs). For the fungal community, differential abundance analysis showed that within the endosphere or rhizosphere most of the significantly differentially abundant groups were reduced in abundance compared to the bulk soil. One taxon however was significantly more abundant in the rhizosphere (*Mortierellaceae*), and one was significantly more abundant in the endosphere (*Parmeliaceae*) (Figure 3.5 B). No significantly differentially abundant archaeal families were detected.

**Table 3.1. SIMPER outputs for field grown wheat**

Similarity percentages breakdown (SIMPER<sup>294</sup>) is a statistical package that scores the percent contribution of each microbial taxa to a community level change. This table shows the percent contribution (% contrib) of each family which significantly (see p-value) contributed to the shift in community composition from the bulk soil to rhizosphere, bulk soil to endosphere, and rhizosphere to endosphere for field grown stem elongation growth phase plants (N=3).

Bulk soil-Rhizosphere			Comparison- Bulk soil-Endosphere		
Family	% contrib *	p-value	Family	% contrib *	p-value
<i>Streptomycetaceae</i>	7.82061	0.10891	<i>Streptomycetaceae</i>	16.39547552	0.00990
<i>Burkholderiaceae</i>	5.59509	0.77227	<i>Burkholderiaceae</i>	6.073676003	0.00990
<i>Firmicutes</i>	4.89642	0.10891	<i>Firmicutes</i>	5.602921016	0.00990
<i>Acidobacteria</i> Subgroup 6	4.74396	0.93069	<i>Acidobacteria</i> Subgroup 6	5.547716313	0.00990
<i>Sphingobacteriaceae</i>	4.08710	0.60396	<i>Sphingobacteriaceae</i>	4.910748857	0.04950
<i>Bacillaceae</i>	3.61463	0.00990	<i>Bacillaceae</i>	4.262550068	0.00990
<i>Methyloligellaceae</i>	3.12084	0.00990	<i>Methyloligellaceae</i>	2.413515939	0.00990
<i>Promicromonosporaceae</i>	3.02557	0.17821	<i>Promicromonosporaceae</i>	2.307694859	0.01980
Solirubrobacterales	2.78842	0.17821	Solirubrobacterales	2.038018274	0.00990
Comparison- Rhizosphere-Endosphere			* % contrib = Percent Contribution		
Family	% contrib *	p-value			
<i>Sphingobacteriaceae</i>	20.89116	0.01980			
<i>Burkholderiaceae</i>	7.499111	0.01980			
<i>Micrococcaceae</i>	4.78363	0.00990			
<i>Acidobacteria</i> Subgroup 6	4.240018	0.35643			
<i>Bacillaceae</i>	4.209923	0.65346			
<i>Pseudomonadaceae</i>	3.371828	0.98019			
<i>Rhizobiaceae</i>	2.840192	0.68316			
<i>Methyloligellaceae</i>	2.130237	0.29703			
<i>Spirosomaceae</i>	2.044043	0.34653			
Solirubrobacterales	1.735366	0.02970			

Quantitative PCR (qPCR) was used to estimate the total abundance of archaeal and bacterial 16S rRNA genes and fungal 18S rRNA genes within each root associated compartment (Figure 3.4). For the bacterial 16S rRNA gene,  $4.98 \times 10^6$  16S rRNA gene copies / 50 ng DNA were detected from the bulk soil, significantly fewer than were detected from the rhizosphere ( $7.03 \times 10^6$  16S rRNA gene copies / 50 ng DNA, Tukey's HSD,  $p < 0.01$ ). The endosphere showed the lowest bacterial 16S rRNA gene copy number ( $1.19 \times 10^6$  16S rRNA gene copies / 50 ng DNA), significantly fewer than for either of the other compartments (Tukey's HSD,  $p < 0.01$  for both comparisons). Fungi outnumbered bacteria and archaea by more than an order of magnitude within the endosphere ( $1.72 \times 10^7$  18S rRNA gene copies / 50 ng DNA) (Figure 3.4). This may indicate that fungi are more abundant within the endosphere but could also be a product of the higher 18S rRNA gene copy number per genome

within some fungi<sup>307</sup>. When comparing bulk soil to the endosphere, archaeal 16S rRNA gene copy number decreased by two orders of magnitude in the endosphere ( $1.18 \times 10^6$  16S rRNA gene copies / 50 ng bulk soil DNA, compared to  $3.89 \times 10^3$  16S rRNA gene copies / 50 ng endosphere DNA), whilst the same comparison for the fungal 18S rRNA gene copy number showed an increase of two orders of magnitude ( $4.32 \times 10^5$  18S rRNA gene copies / 50 ng bulk soil DNA,  $1.72 \times 10^7$  18S rRNA gene copies / 50 ng endosphere DNA). Despite the comparatively lower 16S rRNA gene copy number found in most archaeal genomes<sup>308</sup> the magnitude of the observed difference likely demonstrates archaea colonise the root in much lower numbers than the other root microbiota.

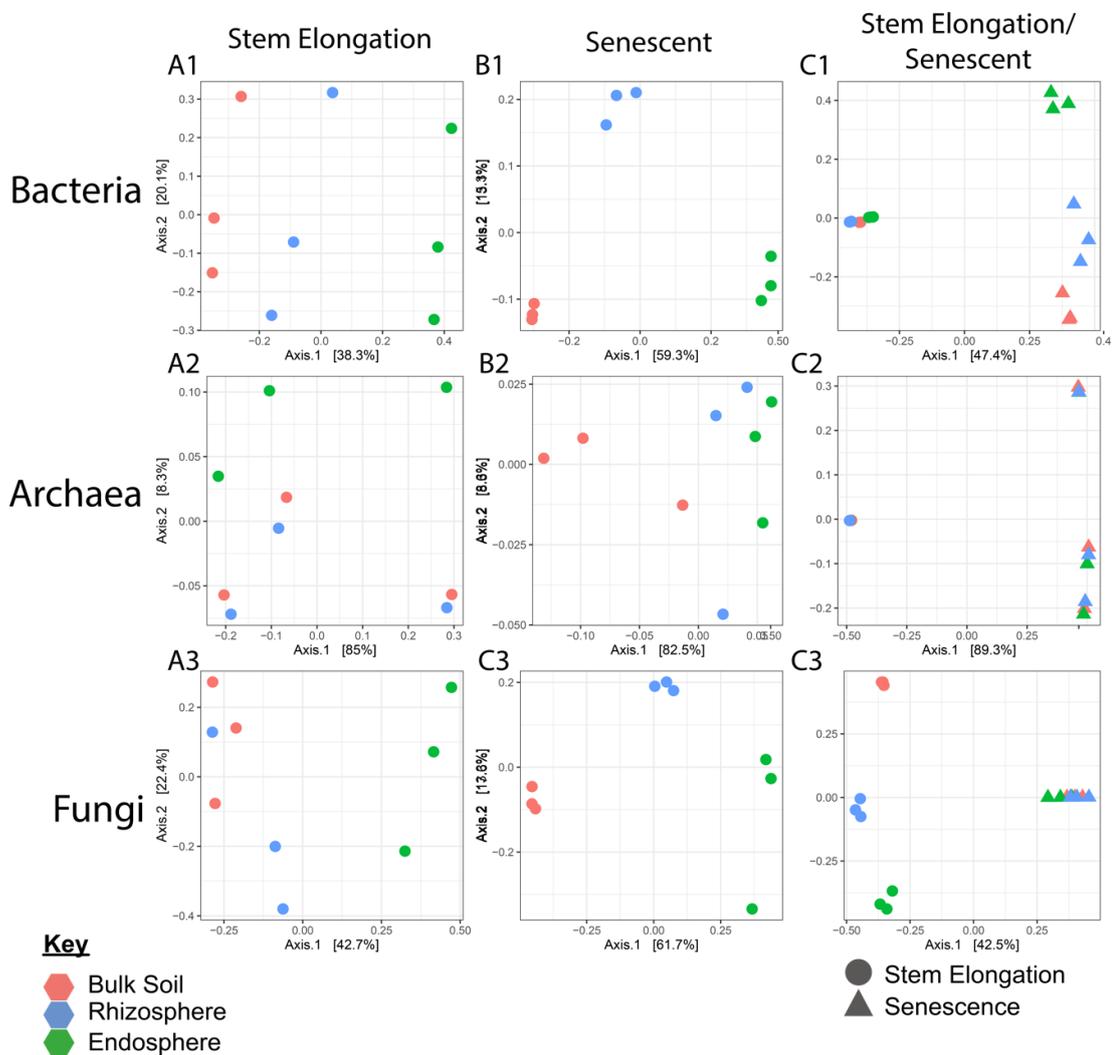


Figure 3.2 Principal Coordinates Analysis (PCoA) performed on Bray Curtis dissimilarities, to assess the similarity of the bacterial, archaeal, or fungal communities associated with the three wheat roots compartments, for plants at the stem elongation growth phase or during senescence. Colours indicate root compartment; green = endosphere, blue = rhizosphere, and pink = bulk soil. N=3 replicate plants per treatment. A1, A2, and A3 show PCoA for Plants cultivated at the Church Farm field studies site at the stem elongation growth phase. B1, B2, and B3 show data from plants sampled from the same site after developmental senescence. C1, C2, and C3 show comparisons between stem elongation growth phase (circles) and senescent plants (triangles).

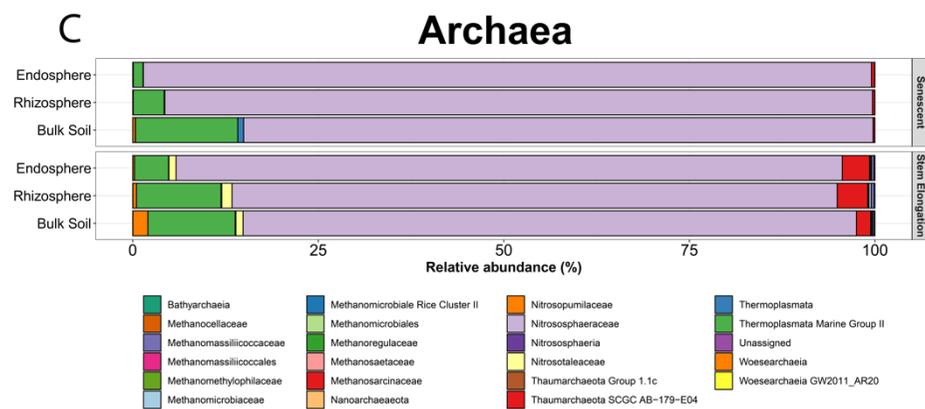
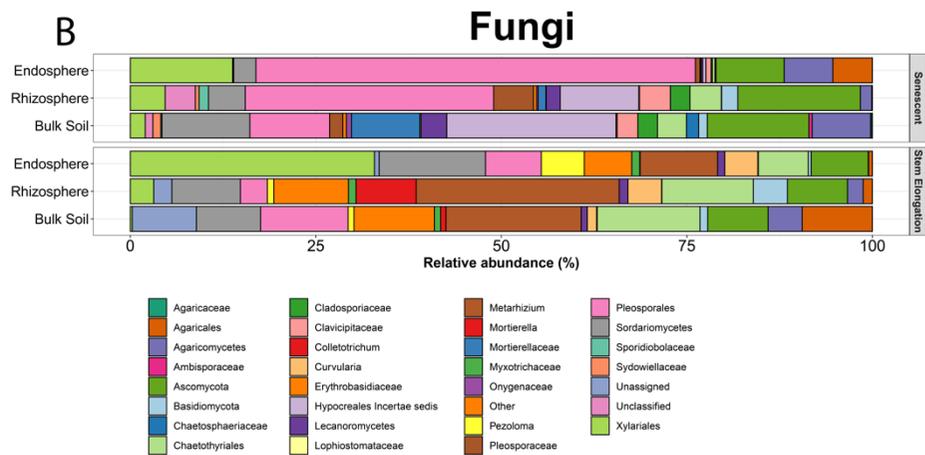
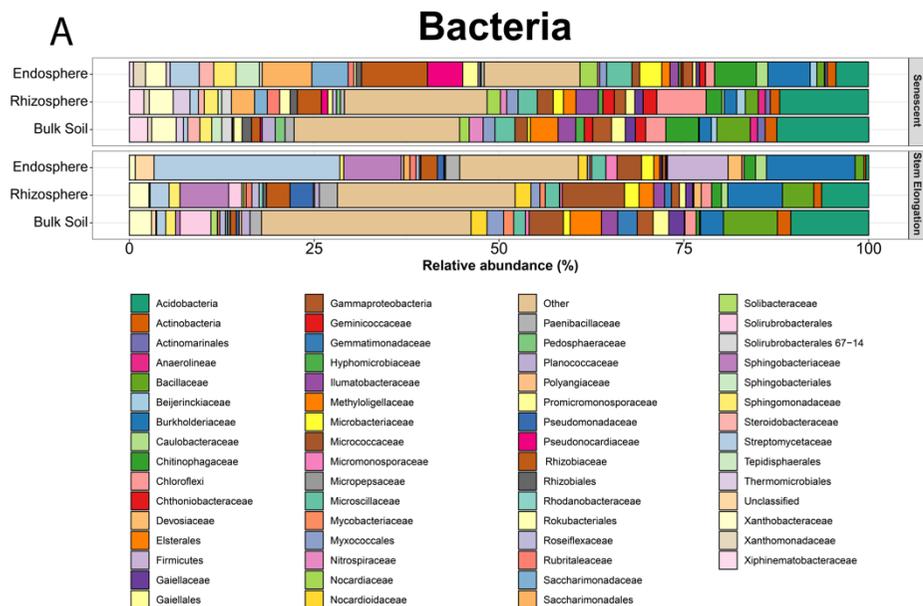


Figure 3.3 Metabarcoding performed to profile the bacterial, archaeal, and fungal communities across three root compartments for senescent and stem elongation growth phase plants. Bars show the mean relative abundance (%) of each bacterial, fungal, or archaeal taxon within the endosphere, rhizosphere, or bulk soil of paragon var. wheat plants sampled at the stem elongation growth phase or during developmental senescence. Plants were grown at the Church Farm field studies site (N=3 replicate plants per condition). Colours indicate different microbial taxa (bacterial, fungal, or archaeal). Within stacked bars, taxa are shown in reverse alphabetical order (left to right). The “Other” category includes all taxa with a median relative abundance of 0.05% or less. ASVs were assigned and are presented to the family level, where family-level taxonomic assignments were unavailable the next highest taxonomic assignment was presented.

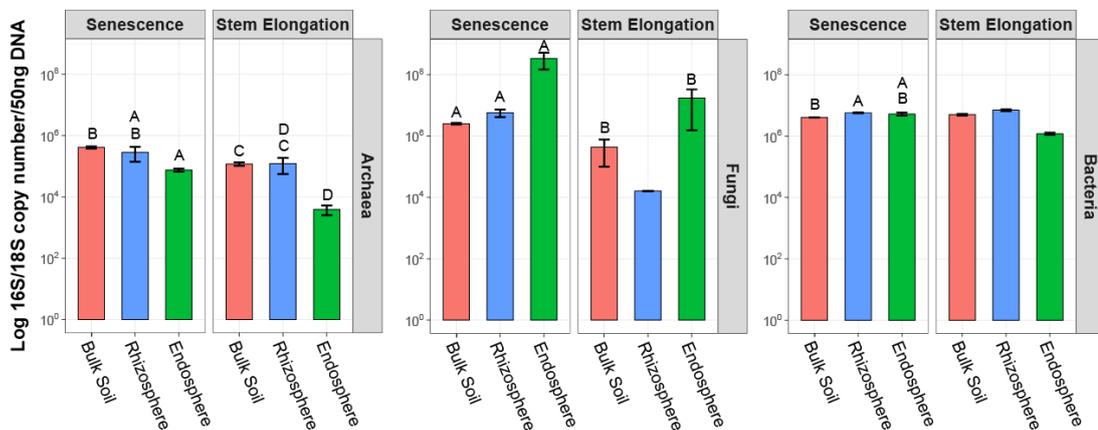


Figure 3.4 qPCR performed on root DNA extract samples to assess the absolute abundance of fungi, bacteria, or archaea within the root microbiome. Bars show the mean log 16S or 18S rRNA gene copy per 50 ng of DNA within the endosphere, rhizosphere, or bulk soil compartment of plants. Plants were sampled from the Church farm field studies site during developmental senescence or during the stem elongation growth phase. N=3 replicate plants per treatment. Bars represent  $\pm$  standard error of the mean. For each group (archaea, bacteria or fungi) a general linear model (GLM) was used to test for a general effect of growth phase or compartment on 16S or 18S rRNA gene abundance. As a post-hoc a Tukey test was ran to identify significant differences between individual conditions. Comparisons were made between each compartment at either senescent or stem elongation growth phases, and individual compartments were compared across growth phases. For all three groups, for samples labelled A, B, C, or D, no significant difference was found (Tukeys HSD,  $p > 0.05$ ), for all other comparisons a significant difference was found (Tukeys HSD,  $p < 0.05$  for all).

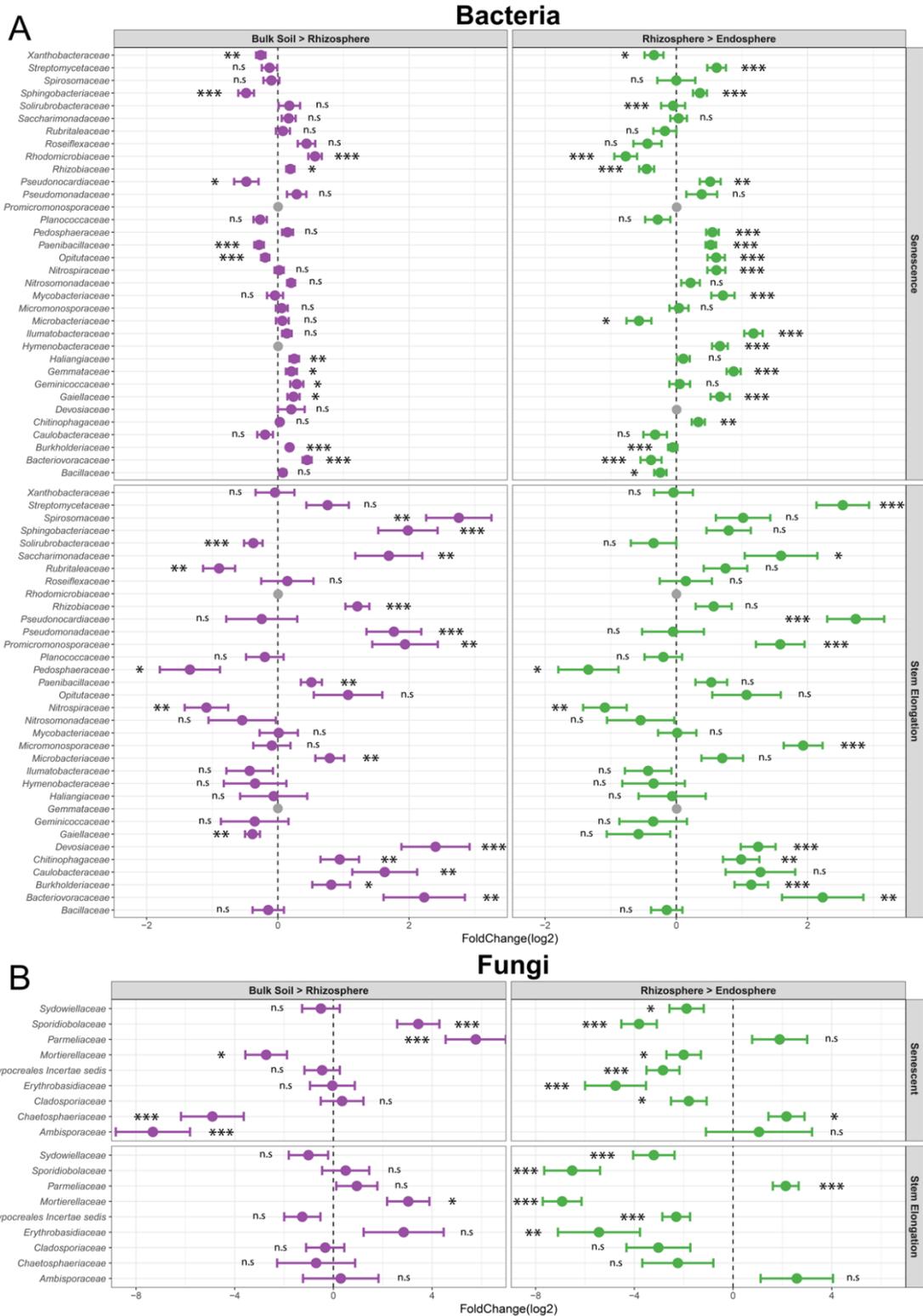


Figure 3.5 . Results of differential abundance analysis to identify bacterial or fungal taxa enriched within the rhizosphere or endosphere of plants sampled during the stem elongation growth phase or during senescence. No significantly differentially abundant archaeal taxa were identified. Dots show the  $\log_2$  fold change of different bacterial or fungal families and error bars show  $\pm$  log fold change standard error. Results are from N=3 replicate plants per treatment. Shown are: **A** Bacterial and **B** fungal families that were differentially abundant between the bulk soil and the rhizosphere, or between the rhizosphere and the endosphere for stem elongation or senesced plants. Grey dots indicate no output was obtained for that family. Analysis was performed using DESeq2. If a family had a base mean > 200 and a significant p-value (significance cut-off  $p < 0.05$ , Bonferroni corrected) in one or more comparisons, \* indicates  $p < 0.05$ , \*\* indicates  $p < 0.01$ , \*\*\* indicates  $p < 0.001$ , and n.s. indicates  $p > 0.05$ . For all complete statistical outputs see Supplementary Tables S.2, S.4 and S.6.

### 3.2 The effect of developmental senescence on the wheat root microbiome

We next aimed to investigate the effect of developmental senescence on the root microbial community and to identify microbial taxa associated with the roots of living plants that displayed reduced relative abundance after developmental senescence. Senescent plants were sampled from the same field site as the plants sampled during stem elongation growth phase. Analysis of rRNA gene copy number (from qPCR experiments, Figure 3.4) showed that plant growth phase significantly influenced the abundance of bacteria (growth phase in a linear model: F-value = 4.86,  $p < 0.05$ ) and archaea (F-value = 10.55,  $p < 0.01$  in a linear model) within the root microbiome. Comparing specific compartments for each group showed that, while there was no significant difference in the abundance of bacteria within the bulk soil or rhizosphere sampled at either growth phase (Tukey's HSD,  $p > 0.05$ ), the abundance of bacteria increased significantly within the endosphere after senescence (Tukey's HSD,  $p < 0.001$ ). The fungal 18S rRNA gene copy number was significantly reduced in the rhizosphere after senescence (Tukey's HSD,  $p < 0.05$ ) but increased by an order of magnitude in the endosphere, although this increase was not statistically significant (Tukey's HSD,  $p > 0.05$ ), likely due to variation across replicates. For archaea there were no statistically significant differences in 16S rRNA gene copy number between the two growth phases for any compartment.

After senescence the most abundant bacterial taxon within the endosphere had changed compared to the stem elongation growth phase. Now, the taxon with the greatest average relative abundance within the endosphere was *Rhizobiaceae* (senescence 8.9%, stem elongation 2.2%), followed by *Saccharimonadales* (senescence 6.7%, stem elongation 0.8%). While greatly reduced, *Burkholderiaceae* were still among the top three most abundant endosphere taxa (senescence 5.7%, stem elongation 12%). *Streptomycetaceae* however were much less abundant in both the endosphere (senescence 3.9%, stem elongation 25.1%) and rhizosphere (senescence 1.1%, stem elongation 2.6%). The most abundant rhizosphere taxa were both annotated to the phyla level and were a greater proportion of the community after senescence, Acidobacteria (senescence 12%, stem elongation 7.4%) and Chloroflexi (senescence 6.7%, stem elongation 1.4%). The third most abundant taxon, *Rhizobiaceae*, showed the same relative abundance at both growth stages (senescence 3.2%, stem elongation 3.2%). *Burkholderiaceae* average relative abundance was much lower in the rhizosphere after senescence (senescence 1.7%, stem elongation 7.4%) (Figure 3.3, A).

Overall, for senescent plants, the most abundant rhizosphere and endosphere bacterial taxa showed a lower average relative abundance than the most abundant taxa during the stem elongation growth phase. This indicated that the root associated community after senescence is more diverse, and this may be a result of weaker selection from the host plant for the maintenance or exclusion of specific taxa. For the fungal community, the most abundant taxon within the endosphere were the order Pleosporales (59.2%), followed by the Xylariales (13.7%), and the Ascomycota (9.2%). Pleosporales were also the most abundant in the rhizosphere (33.5%), followed by Ascomycota (16.5%), and *Hypocreales Incarte Sedis* (10.6%)

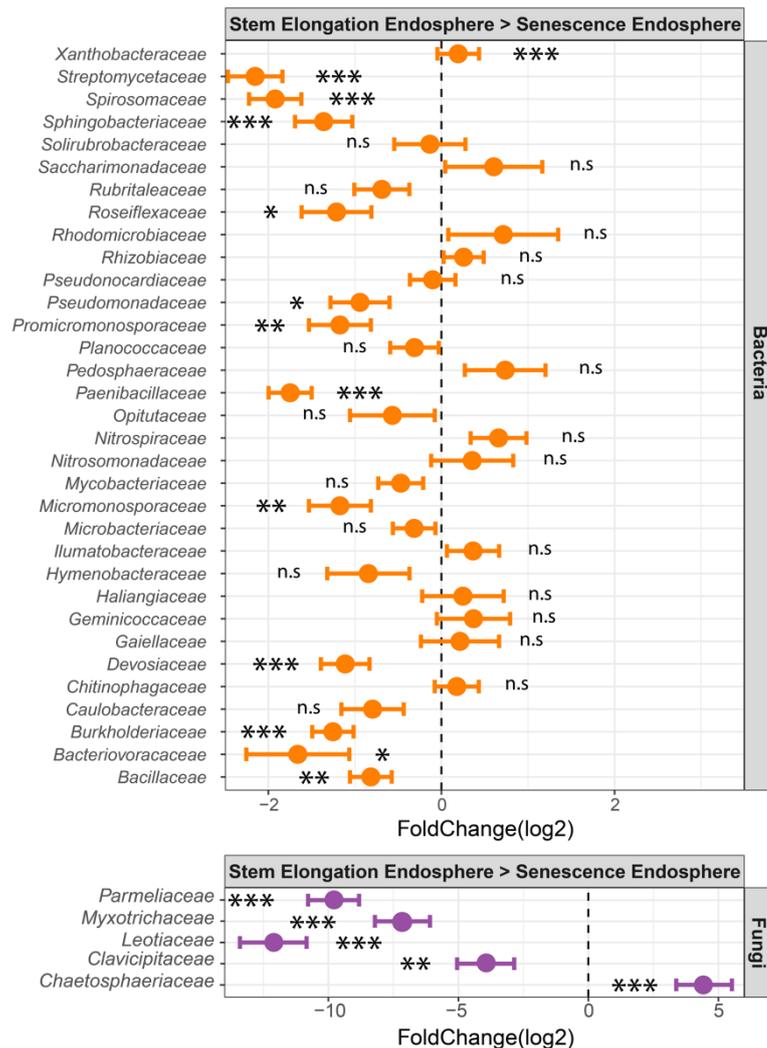


Figure 3.6 Differential abundance analysis to identify bacterial or fungal taxa significantly increased or reduced in abundance within the endosphere of senesced plants compared to stem elongation growth phase plants. No significantly differentially abundant archaeal taxa were detected. Dots show the log<sub>2</sub> fold change of different bacterial or fungal families and error bars show ± log fold change standard error. Results are from N=3 replicate plants per treatment. Results show bacterial and fungal taxa that were differentially abundant between the endosphere of stem elongation growth phase plants and senesced plants. Analysis was performed using DESeq2. If a family had a base mean > 200 and a significant p-value (significance cut-off p < 0.05, Bonferroni corrected) in one or more comparison, data for that taxon was plotted for all comparisons, \* indicates p < 0.05, \*\* indicates p < 0.01, \*\*\* indicates p < 0.001, and n.s. indicates p > 0.05. For all complete statistical outputs see Supplementary Tables S.5 and S.7.

(Figure 3.3, B). While the dominance of Pleosporales might imply a much simpler fungal community, this order of fungi contains a vast diversity of fungal species and so may include a broad range of fungal rhizosphere functions, including those of numerous well-known cereal crop pathogens, which could benefit from a deficient host immune response during developmental senescence. The archaeal community after senescence was almost entirely comprised of the AOA family *Nitrososphaeraceae* (98.1% endosphere, 95.3% rhizosphere) (Figure 3.3, C).

Both fungal and bacterial community composition differed significantly across the three different root compartments of senescent plants, as clearly demonstrated by PCoA (Figure 3.2, B) and permanova for all three microbial groups (Table 3.2). In addition to this, PCoA showed a clear difference between the microbial communities associated with senescent or stem elongation growth phase plants, however, they also indicated that the root community was much more variable for senescent plants compared to those in the stem elongation phase (Figure 3.2, C). Permanova analysis corroborates this observation as, whilst this showed a significant effect of plant growth phase on overall community composition for all three microbial groups (permanova, bacterial:  $R^2=0.47$ ,  $p < 0.001$ , archaeal:  $R^2=0.89$ ,  $p < 0.001$ , fungal:  $R^2=0.42$ ,  $p < 0.001$ ), betadisper analysis indicated that microbial community dispersion was not equal between the two growth phases ( $p < 0.01$  for all), i.e. the senescent growth phase showed greater community variability compared to the stem elongation phase.

**Table 3.2 Permanova results, senescent/stem elongation**

Permanova (permutational analysis of variance) comparing the bacterial, archaeal, or fungal community composition within senescent plants to that of stem elongation growth phase plants. This showed a significant shift in community composition for all three (see p-values) and showed that developmental stage explained the majority of variance (see  $R^2$  values). All tests were run with 999 permutations. N=6 for all comparisons (rhizosphere and endosphere communities).

<b>Test for effect of root compartment on senescent communities</b>			
<b>Community</b>	<b>Permutations</b>	<b>R<sup>2</sup></b>	<b>p-value</b>
Archaea	999	0.68	0.01
Bacteria	999	0.74	0.005
Fungi	999	0.73	0.005

For individual taxa, differential abundance analysis showed that 16 bacterial and fungal taxa were significantly less abundant within the endosphere of senesced plants compared to at the stem elongation growth phase ( $p < 0.05$ , Figure 3.6). The largest change in abundance was a two-fold reduction in the family *Streptomycetaceae*, and there was also a significant reduction in the relative abundance of the families *Burkholderiaceae* and *Sphingobacteriaceae* in senescent plants (Figure 3.6). This implies that these taxa may require input from the living plant in order to persist within the endosphere. No archaeal taxa demonstrated significant changes in their abundance across root compartments or between growth phases, and the root archaeal community was consistently dominated by the AOA family *Nitrososphaeraceae*. For the fungal community, differential abundance analysis indicated that the abundance of most taxa was significantly reduced in senescent plants, except for *Chaetosphaeriaceae*, which showed a four-fold increase during senescence when compared to the stem elongation phase.

### **3.3 Archaea within the wheat root microbiome**

#### **3.3.1 The efficacy of short-read amplicon sequencing for archaeal diversity studies**

As shown in Figure 3.3 and in 4.2, the archaeal community was dominated by two families of AOA (*Nitrososphaeraceae* and *Nitrosotaleaceae*), which were abundant in all root compartments. *Nitrosotaleaceae* dominated in the more acidic Levington F2 compost whereas *Nitrososphaeraceae* was most abundant in the neutral pH Church Farm soil (Figure 4.2). Both developmental senescence and soil type were found to significantly influence archaeal community composition, though no selection of specific archaeal lineages within the endosphere was detected by SIMPER or differential abundance analysis, and PCoA did not show a strong effect of compartment on community composition (Figure 3.2, Chapter 4 Figure 4.1). Contrary to this, there was a small but significant shift in the archaeal community composition overall across compartments (archaeal permanova:  $R^2=0.86$ ,  $p = 0.001$ ), and a betadisper analysis was not significant ( $p > 0.01$ ), demonstrating this was not due to difference in dispersion between compartments. From these findings it is unclear whether there was any major selection of archaeal taxa by the wheat roots. Denaturing gradient gel electrophoresis (DGGE) analysis performed on both the archaeal 16S rRNA and *amoA* genes showed a clear shift in the archaeal community across compartments (Figure 3.7), supporting the hypothesis that there

is in fact selection for specific archaeal lineages across root compartments. Unfortunately the archaeal 16S rRNA gene database lacks the established framework of its bacterial counterpart<sup>251</sup>, i.e. archaeal sequence databases do not enjoy an established and widely accepted phylogeny through which sequences can be assigned an identity. This, coupled with the lack of known diversity or strain characterisation within many archaeal taxa, makes it difficult to achieve good taxonomic resolution from short read amplicon sequencing of the archaeal 16S rRNA gene. We hypothesised therefore that this discrepancy between DGGE and amplicon sequencing arose from the lack of detailed taxonomic representation within the database used to analyse the sequencing data. Despite these limitations, this study has revealed that AOA dominate the archaeal community associated with wheat roots regardless of soil type, and that the abundance of archaea within the root is highest in agricultural soil and increases later in the life cycle of the plant. There may be selection of specific archaeal lineages across the root compartments, as indicated by DGGE and statistical analysis, but further detailed analysis is required.

### **3.3.2 Assessing plant-growth promoting potential of an ammonia oxidising archaeon**

Given the observation that families of AOA dominate the root associated archaeal community, and that there is some evidence for AOA plant growth promotion<sup>236</sup>, the efficacy of AOA to promote the growth of *T. aestivum* was assessed. Plants were given one of three treatments; a concentrated inoculum of AOA strain *Nitrososocosmicus franklandus* C13 suspended in MSK plant growth medium (AOA), and either an uninoculated plant growth medium control (medium treatment), or a dH<sub>2</sub>O control (dH<sub>2</sub>O). MSK plant growth medium was used to suspend archaeal cells as, contrary to data presented by Song *et al.*<sup>236</sup>, archaeal growth medium (Chapter 2, Table 2.1) was found to inhibit plant growth in an early experiment, likely due to high salt levels (Supplementary Figure S.1). An enrichment culture experiment showed that *N. franklandus* C13 is able to grow in MSK medium (Supplementary Figure S.2), meaning the solution can be used for the re-suspension of harvested *N. franklandus* C13 cells without jeopardising their viability. The strain, *N. franklandus* C13, was chosen for its close relatedness to the strain used by Song *et al.*<sup>236</sup>, though there are some key differences between the strains growth characteristics, most notably optimal temperature<sup>205,236,309</sup>. Whilst there was a significant effect of treatment on total dry weights (Kruskal-Wallis, chi-squared = 15.7,  $p < 0.001$ ), a Dunns test showed that there was no significant difference

between plants treated with AOA when compared to those treated with the uninoculated medium (Table 3.3). Comparing the medium treatment to the medium + AOA treatment would suggest that the addition of AOA to MSK medium negated the plant growth promoting effects of the nutrient solution, possibly due to oxidation of ammonia by *N. franklandus* C13. Thus, the present experiment shows that the AOA strain *N. franklandus* C13 did not promote plant growth.

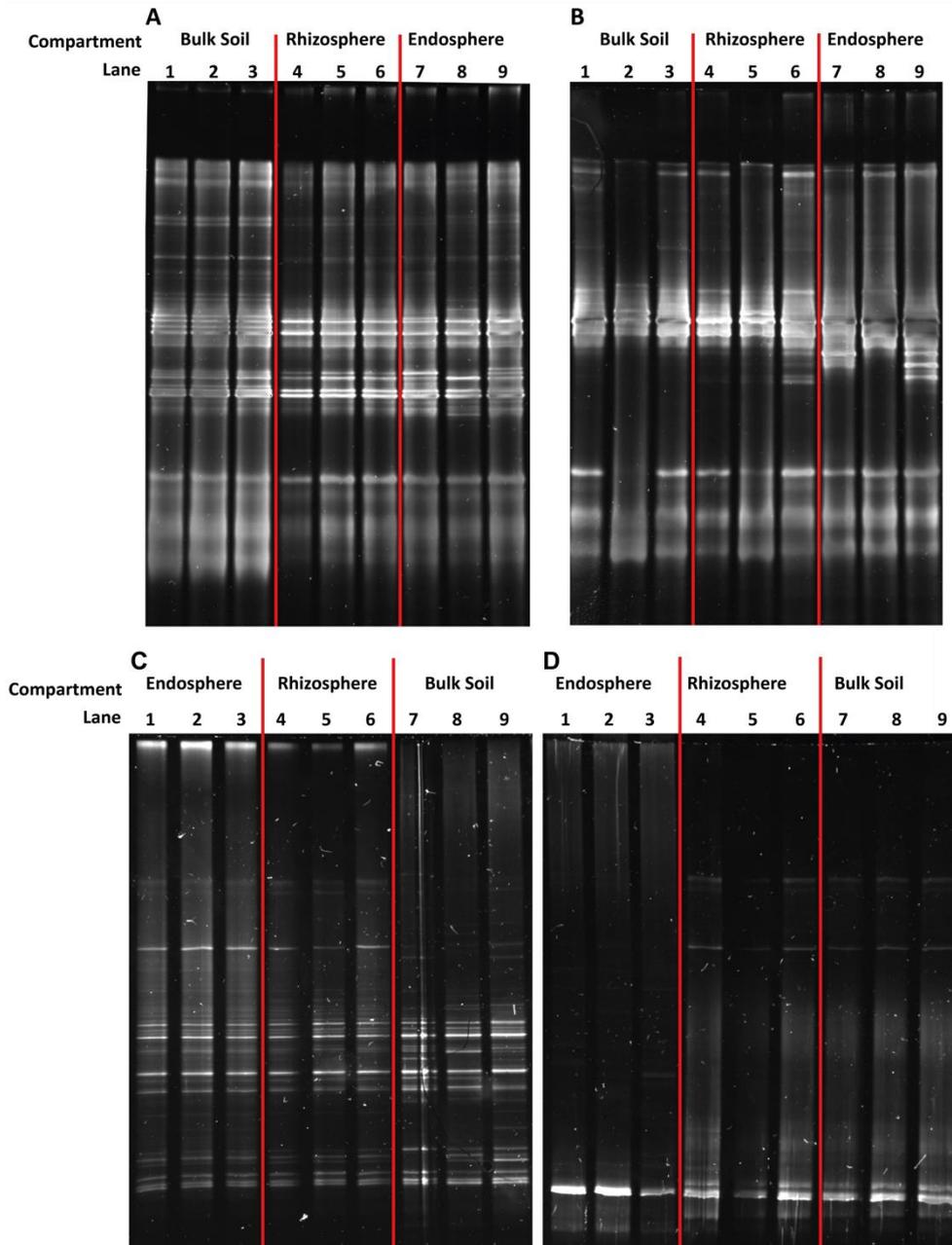


Figure 3.7 Denaturing gradient gel electrophoresis (DGGE) performed to assess archaeal 16S rRNA gene (**A** and **B**) or *amoA* (**C** and **D**) diversity across the three root compartments. Columns show DGGE on amplicons from the bulk soil, rhizosphere or endosphere of wheat grown under laboratory conditions in agricultural soil (**A** and **C**) or Levington F2 compost (**B** and **D**) (N=3). Primers are indicated in the Table 2.4.

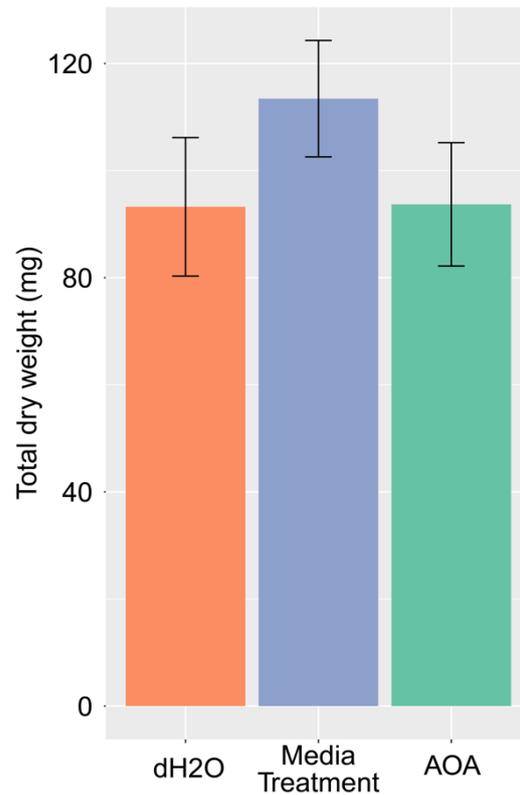


Figure 3.8 Plant growth promotion experiment assessing the ability of ammonia oxidising archaeon strain *Nitrosocosmicus franklandus* C13 to promote the growth of wheat in agricultural soil. Bars show the mean total plant dry weight (mg) four weeks after plants were treated with AOA strain *N. franklandus* C13 (green), and either a media (purple) or dH<sub>2</sub>O (orange) control. N=36 plants across three experimental replicates. Error bars show  $\pm$  SEM.

**Table 3.3. Dunns test results**

Dunns test to identify significant differences in the total mean dry weights for plants treated with ammonia oxidising archaeon *Nitrosocosmicus franklandus* C13, blank medium, or dH<sub>2</sub>O.

Dry Weight Comparison	p-value
AOA - dH <sub>2</sub> O	0.0028 *
AOA - Media	0.1323
Media - dH <sub>2</sub> O	0.0001 *

It is unclear why the present study does not corroborate the plant growth promoting effects of AOA observed by Song *et al.*, there are four possible explanations. (1) The PGP effect observed is not a broadly held phenotype amongst AOA, and the strain used by Song *et al.* (*Nitrosocosmicus oleophilus* MY3<sup>236,309</sup>) may be rare in this regard. (2) Song *et al.* conducted the AOA PGP experiment under axenic conditions, the effect may not be reproducible in a less controlled but more agriculturally comparable soil environment, such as used by the present work. (3) *N. franklandus* C13 may not be active at the ambient temperatures. Most PGP experiments are performed under controlled laboratory or greenhouse conditions between 21°C and 25°C, and the present work was performed at 21°C. Whilst the strains optimum temperature is 30°C, at ambient temperatures *N. oleophilus* MY3 is active<sup>309</sup>, whereas the optimum temperature for *N. franklandus* C13 is 37°C<sup>205</sup>. It must be noted that *N. franklandus* C13 was originally isolated from soil sampled from an agricultural site in (Aberdeenshire, UK)<sup>205</sup> where average yearly temperatures range from 5°C to 14°C. Presumably then *N. franklandus* C13 would be able to function at lower temperatures, though this capability may have been lost through the enrichment culture process. Regardless, these discrepancies in strain temperature preference may be responsible for the contrasting PGP activity of these strains within laboratory experiments. (4) The PGP effect of AOA may be specific to *Arapidopsis* or other *Brassica* species and may not translate to grass species such as UK elite spring bread wheat. It is unclear which of these hypotheses explains the discrepancy between the present results and those of Song *et al.*, and further work would be required to fully understand if *N. franklandus* C13 can promote the growth of wheat, or if other AOA species can promote the growth of wheat in an agriculturally relevant soil system. The present work demonstrates clearly however that under ambient conditions *N. franklandus* C13 is unable to promote the growth of wheat when cultivated in agricultural soil.

### 3.4 Discussion and conclusions

16S rRNA gene & ITS2 region metabarcoding has shown that *T. aestivum* var. Paragon harbour endosphere and rhizosphere communities that are distinct from bulk soil. This shift was most pronounced for the bacterial community, with PCoA showing a clear distinction between the rhizosphere and endosphere. Concurrent with previous observations<sup>68,97</sup>, the rhizosphere effect observed for this wheat cultivar was weak; while at the community level there was a significant shift in the

community from bulk soil to rhizosphere, the same main microbial taxa are present, with only minor, but consistent, changes in their relative abundance (Figure 3.3). The most abundant bacterial taxa within the endosphere were the families *Streptomycetaceae* and *Burkholderiaceae*. This might imply a beneficial relationship between these families and the host, given that species from both taxa have plant-beneficial traits<sup>25,37,41,42,214</sup>, and members of the *Streptomycetaceae* family are the active agents in the biocontrol formulations Actinovate and Mycostop<sup>310</sup>. However, species within these groups, such as *Streptomyces scabies*<sup>311,312</sup> or *Burkholderia pseudomallei*<sup>313</sup>, are known as plant pathogens. Given that healthy plants were sampled from the field for this experiment, and pathogenicity is restricted to a small number of closely related *Streptomyces* species (a small minority of the hundreds of species characterised for this genus<sup>314</sup>) it is more likely that the bacteria within the roots in this case are not pathogenic.

Most fungal taxa were reduced in abundance within the endosphere of stem elongation growth phase plants; the one exception to this was the family *Parmeliaceae* which showed a two-fold increase in the endosphere compared to the rhizosphere (Figure 3.5). This group however showed a very low relative abundance overall and was below the 0.05% median relative abundance threshold used to define the “Other” category in Figure 3.3. Given that this family is broadly associated with lichen formation<sup>315</sup> it is difficult to postulate a role for this group within the wheat endosphere, and given its low relative abundance (< 0.05%) it is possible their apparent enrichment was a false positive, resulting from for example a sequencing artifact<sup>316</sup> or overestimation by differential abundance analysis software DESeq2. The most abundant fungi within the endosphere were broad groups such as the Xylariales and their parent class the Sordariomycetes. None of these groups were identified as enriched in the endosphere or rhizosphere by differential abundance analysis, so we cannot know if these unknown fungi are beneficial for the host or are passive community members. Furthermore, the diversity within these groups could include beneficial fungi such as *Trichoderma harzianum*<sup>317</sup>, or root-pathogens such as the take-all fungus (*Gaeumannomyces graminis* var. *tritici*)<sup>318</sup>, both of which fall within the Sordariomycetes, though given that healthy plants were sampled the presence of virulent *G. graminis* is unlikely.

The Xylariales order contains a variety of different fungi, including fungal endophytes<sup>319</sup> and detritivores<sup>320</sup>. While we can identify community level changes in the microbial community, due to the short-read nature of the present sequencing

data, and subsequent high-level taxonomic classifications, it is difficult ascertain fungal functions within the root. For Xylariales it could be hypothesised that are fungal detritivores, attracted to decomposing border like cells (BLCs) and other discarded root tissue. BLCs are detached lateral root cap cells (LRCCs); in actively growing roots during a process called root cap detachment LRCCs are ejected from the root apical meristem to maintain root cap size as the plant grows. These then form BLCs, a covering of loosely associated living cells surrounding the root tips which can persist for days to weeks before decomposition<sup>321</sup>. While alive, BLCs have an important role in root health; they provide protection against fungal pathogens by acting as a physical barrier, and via a range of secretory and signalling activities<sup>188,321</sup>, including the self-destructive secretion of nucleic acid extracellular traps (NETs) which trap and kill invading pathogens but also kill the BLC<sup>322</sup>. BLCs also influence root microbiome composition by contributing to root exudation<sup>321</sup>. These shed BLCs however eventually lose viability and die, additionally other cells within the root apical meristem, such as lateral root cells, regularly undergo programmed cell death<sup>323</sup>. Constant cell-shedding, combined with self-destructive defensive mechanisms, and cell turnover in the actively growing roots generates an environment rich in dead plant matter. After these cells die, they could feasibly form a niche for Xylariales order detritivores; Xylariales decreased in abundance after senescence (senescent 13.7%, stem elongation 32.9%) (Figure 3.3 B) which could indicate that the niche occupied by this fungus is disrupted when root cap detachment is arrested after senescence and when roots are no longer actively growing. Additionally, Xylariales were less abundant in four-week-old lab cultivated plants than in mature field grown wheat (16.5%) (Chapter 4, Figure 4.2). As younger plants will have grown less, and thus shed fewer LRCCs, it could be hypothesised that they have had less time to enrich for Xylariales detritivores compared to mature plants which had been growing and undergoing root cap detachment for ~28 weeks. This supports the hypothesis that detritivorous Xylariales order fungi could be attracted to the root environment to feed on discarded cell matter, and that within the present dataset the Xylariales order represents detritivores that occupy a niche in actively growing plants feeding on discarded root matter, that they increase in abundance as the plant ages, and finally are reduced in abundance after senescence when root shedding activities stop. To conclusively investigate this hypothesis more detailed analysis is required, which could use in vitro experiments assessing the ability of fungal detritivores to colonise BLCs. Alternately ITS2 community profiling different root segments (such as the apical meristem, elongation zone, and differentiation zone) could assess the

quantities of Xylariales order detritivores within root segments with different levels of cell turnover.

For the archaeal community, AOA were found to dominate in all root compartments. Whilst no selection of specific archaeal lineages within the root could be detected via sequencing, DGGE did indicate a possible shift in community composition across root compartments. The potential for interactions between soil AOA and plant roots remains largely unexplored. There is some limited evidence however which may indicate an influence of terrestrial plant root exudates on archaeal communities <sup>255</sup>, and whilst the present work found no clear evidence that the total abundance of archaea changed within the rhizosphere, one study observed a negative correlation between archaeal abundance and plant root exudates <sup>168</sup>. There is also some evidence that AOA can promote plant growth <sup>236</sup>. The nature of these interactions however still remains unclear. There is now mounting evidence that archaeal communities are influenced by plants or plant-derived metabolites within the soil, even if they do not utilise host derived carbon. In the future, longer read methods or metagenomics could be applied to better investigate archaeal community dynamics within the root microbiome.

At the onset of developmental senescence there were significant changes in the abundance of numerous bacterial and fungal taxa. To our knowledge, the wheat root community has not previously been assessed after senescence, though development has been shown to significantly alter the wheat rhizosphere community <sup>95,96</sup>. Intriguingly, despite the fact *Streptomycetaceae* constituted on average 25% of the bacterial endosphere community during the stem elongation growth phase, after the plants senesced, this was dramatically reduced. For senescent plants *Streptomycetaceae* constituted on average just 3.9% of the endosphere community, and differential abundance analysis showed a two-fold reduction in the family's abundance (Figure 3.6). This a surprising result for a bacterial group typically associated with the breakdown of dead organic matter within soils <sup>312</sup>. As plant tissues senesce and die a process of ecological succession occurs, where the tissues are colonised by different microbes (particularly fungi) successively as different resources within the plant tissues are degraded <sup>324,325</sup>. The first microorganisms to colonise will be those rapidly metabolising sugars and lipids, followed later by more specialist organisms which will breakdown complex molecules like lignin and cellulose. While these later stages are typically attributed to fungi, *Streptomycetaceae* are known to degrade complex plant derived molecules such as hemicellulose and insoluble lignin <sup>312,326</sup>. It could be that the sampling

timepoint (late in the developmental senescence process, but prior to most biomass degradation) was too early in this succession process for any biomass fuelled *Streptomyetaceae* proliferation to be obvious. This however cannot explain the reduced abundance of *Streptomyetaceae* in senesced roots compared to the actively growing plants. One hypothesis could be that this is caused by a lack of active input from the plant. As the host senesces and resources are diverted to the developing grain <sup>223</sup>, and as root growth, root cap detachment, and root cell turnover is arrested, host derived resources may no longer be available to support *Streptomyetaceae* growth in the endosphere. Numerous other bacterial taxa were less abundant during developmental senescence, such as *Burkholderiaceae*, *Pseudomonadaceae*, and *Sphingobacteriaceae*. It may be that changes in root exudation profiles caused a reduction in the relative abundance of these bacteria, as they are no longer supported by host derived nutrients.

For the fungal community one family, *Chaetosphaeriaceae*, was significantly enriched within the endosphere as the plant senesced. This family represents a relatively diverse group of fungi, although members of this group such as *Chaetosphaeria* are known to reproduce within decomposing plant tissues, which may explain the families four-fold increase in abundance after senescence <sup>327</sup>. In terms of the overall fungal community composition (Figure 3.3, B), the greatest change during senescence was in the Pleosporales group which dominated the endosphere and rhizosphere. This colonisation by Pleosporales may have contributed to the observed increase in fungal abundance during senescence. This group was excluded from the differential abundance analysis which focused on lower taxonomic ranks. Pleosporales is an order of fungi containing over 28 families <sup>328</sup>, and, as discussed for other fungal groups, such a high diversity makes the ecological role of this group difficult to postulate. Some families within the Pleosporales are associated with endophytic plant parasites <sup>328</sup>, including necrotrophic pathogens of wheat *Pyrenophora tritici-repentis* and *Parastagonospora nodorum* <sup>329</sup>. Necrotrophic pathogens specialise in colonising and degrading dead plant cells, and senescent tissues are thought to provide a favourable environment for necrotrophs <sup>225</sup>. It is interesting to note that this increased fungal colonisation correlated with reduced abundance of fungi-suppressive endophytic bacteria such as *Streptomyetaceae* <sup>41,42</sup> and *Burkholderiaceae* <sup>37</sup> during developmental senescence. This leads us to hypothesize that as the plant senesces, and root developmental activity and small metabolite exudation is arrested, that the plant is no longer able to maintain populations of core anti-fungal endosphere and

rhizosphere bacteria such as *Streptomycetaceae*, *Burkholderiaceae*, and *Pseudomonadaceae*. As a result, increased fungal root colonisation is seen as Pleosporales order fungi colonise the root and being to breakdown the root tissue. In the future root microscopy, using FISH or synthetic communities of fluorescently labelled core root associated bacteria, could be used to characterise interactions between core microbiota and parasitic or necrotrophic fungi within the root as plants senesce. Alternately, a sequencing-based approach could profile microbial community dynamics using longer reads or more timepoints to track the microbial dynamics of this process in more detail.

In conclusion, this chapter has profiled the bacterial, archaeal, and fungal communities associated with Paragon var. *T. aestivum* roots. This has shown that the endosphere harbours a unique community compared to the rhizosphere and bulk soil. Differential abundance analysis showed that a number of bacterial taxa, and one fungal taxon, were enriched in the endosphere or the rhizosphere. Most notable were the families *Streptomycetaceae* and *Burkholderiaceae*, which were the taxa showing the greatest relative abundance within the endosphere. After developmental senescence these taxa also showed the greatest abundance-reduction; the most-reduced taxon after abundance was *Streptomycetaceae*, which was reduced two-fold after senescence. Study of the fungal and archaeal communities was difficult as short read amplicon sequencing could not achieve a high level of taxonomic resolution, though this did reveal that AOA are present inside the root. DGGE analysis of the archaeal community indicated that there may be some selection of specific archaeal lineages across root compartments, but that the limited nature of archaeal sequencing databases truncated our ability to observe these changes using short read amplicon sequencing.

## **Chapter 4. Drivers of microbial diversity within wheat roots**

Sections 4.1 and 4.2 from this chapter have been published previously <sup>198</sup>.

### **Aims**

As with many plants, wheat root associated microbial communities can be influenced by numerous biotic and abiotic factors including crop cultivation practices <sup>63</sup>, climatic conditions <sup>330</sup>, host genetics <sup>97,108,110,111</sup>, and soil parameters <sup>55,84</sup>. In particular, microbial communities and their functions can differ dramatically between different soils and, as a consequence, soil parameters play a central role in shaping the microbial communities associated with plants. To explore the extent to which these factors influence the rhizosphere and endosphere community this chapter aimed to profile the microbial community in different soil types and associated with different varieties of wheat. It is widely accepted that root microbial communities are acquired horizontally from the soil <sup>52</sup>. To show that this is also the case for wheat, the microbial communities associated with wheat seeds were profiled to ascertain if any microbial taxa are present within the seeds that are also present within the roots. Short read Illumina sequencing, while useful for community-level analysis, yields limited taxonomic resolution. To identify specific species and genera within the root compartments this chapter also aimed to use long read amplicon sequencing to recover full length 16S rRNA gene sequences from the root.

### **Results**

#### **4.1 Pot grown wheat as an agriculturally relevant model**

To test whether the microbiome associated with laboratory-grown plants was comparable to that of field grown wheat, plants were grown for four weeks under laboratory conditions in soil collected from the Church Farm site, and the composition of the root microbiome was profiled using 16S rRNA gene and ITS2 metabarcoding. Laboratory-grown plants were sampled during root growth phase, whereas field plants were sampled during the late stem elongation growth phase, meaning laboratory-grown plants were sampled much earlier in the life cycle. PCoA indicated a shift in the endosphere community when comparing field to pot grown wheat (Figure 4.1). The same major microbial families were present within the endosphere of both groups of plants however (Figure 3.2, Figure 4.2).

Whilst statistical analysis did indicate a significant community-level difference between the bacterial and fungal communities respectively (permanova, bacterial:  $R^2 = 0.12$ ,  $p < 0.001$ , fungal:  $R^2 = 0.13$ ,  $p < 0.01$ , archaeal:  $R^2 = 0.13$ ,  $p > 0.05$ ), subsequent pairwise analysis found no significant difference between any specific compartments (Table 4.1). qPCR indicated that the overall abundance of bacteria and archaea was significantly different between the two groups of plants ( $p < 0.05$  in linear models for both microbial groups). There were significantly more archaea within the bulk soil associated with pot-grown plants (Tukey's HSD,  $p < 0.01$ ), however post-hoc analysis did not show a significant difference in the abundance of either archaea or bacteria in the root associated compartments between the different groups of plants (Tukey HSD,  $p > 0.05$  for all). A significantly greater quantity of fungi was detected within the rhizosphere of laboratory-grown plants (Tukey's HSD,  $p < 0.05$ ) and lower quantities of all groups were observed within the endosphere (Figure 4.). Overall, this analysis showed that there is likely a lower microbial abundance within the endosphere of laboratory-grown root growth phase plants when compared to stem elongation growth phase plants cultivated in the field, but that any effects on community composition were subtle and mostly restricted to low abundance taxa, which can sometimes be the result of sequencing artefacts<sup>105</sup>, and due to their low abundance are more vulnerable to stochastic change. As bacterial, fungal, and archaeal communities contained the same major taxa within the endosphere, it was concluded that laboratory-grown plants could serve as an approximate experimental analogue for agriculturally cultivated wheat plants when studying the composition of the root microbial community.

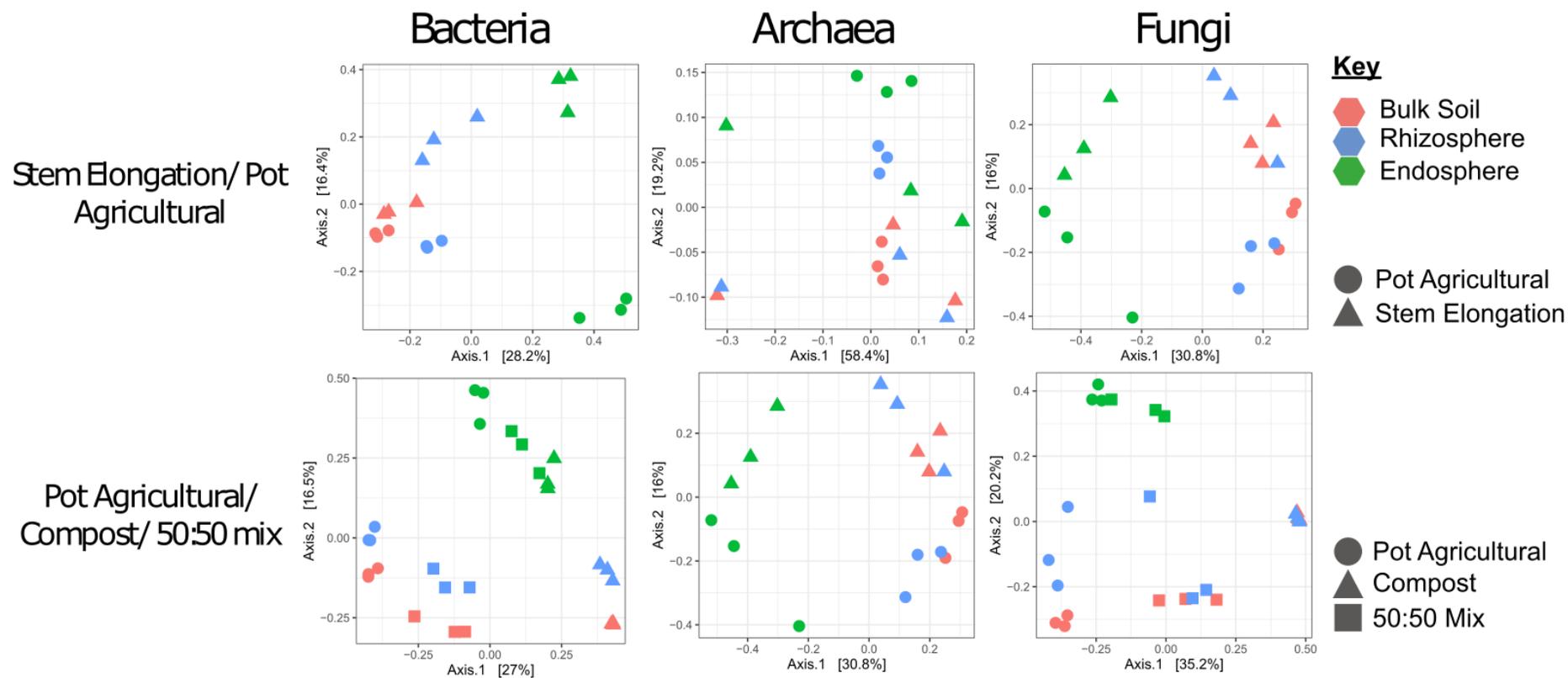
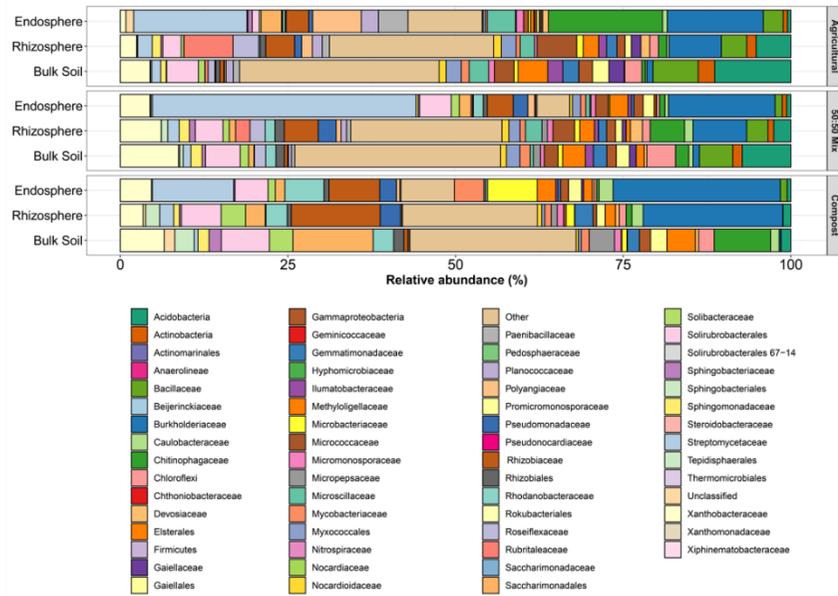
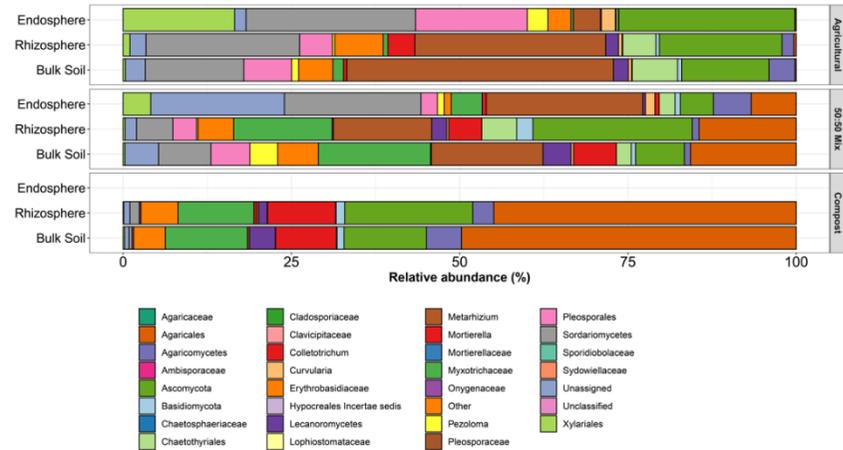


Figure 4.1 Principal Coordinates Analysis (PCoA) performed on Bray Curtis, to assess the similarity of the bacterial, archaeal, or fungal communities associated with the three wheat roots compartments. This was performed to compare the community in mature field cultivated plants at the stem elongation growth phase with four-week-old plants in the root outgrowth stage cultivated in the same soil within the lab (top row). Additionally, this was performed to compare community composition across three different soil types (bottom row). Colours indicate root compartment; green = endosphere, blue = rhizosphere, and pink = bulk soil. N=3 replicate plants per treatment. Comparisons shown are either between 4-week-old laboratory cultivated plants (circles) and stem elongation growth phase field cultivated plants (triangles) (top row) or between communities associated with plants cultivated under laboratory conditions in agricultural soil (circles), Levington F2 compost (triangles), or a 50:50 mix of the two (squares) (bottom row). For the fungal community no data could be acquired for the endosphere of plants cultivated in Levingtons F2 compost. For Levingtons F2 compost only two replicates could be retrieved for the archaeal endosphere community.

## Bacteria



## Fungi



## Archaea

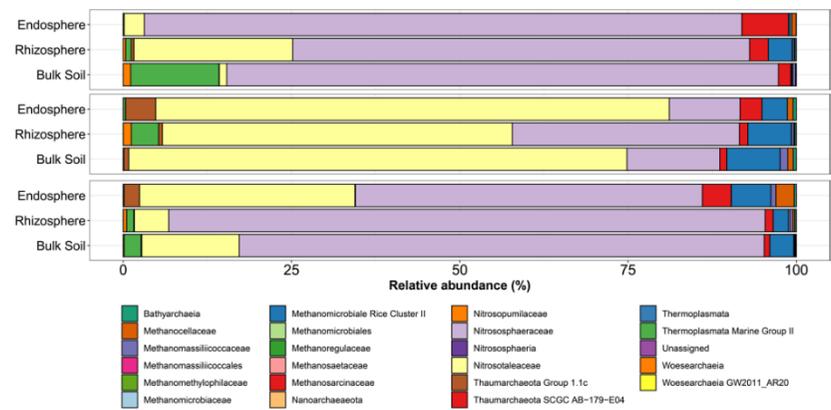


Figure 4.2 Metabarcoding performed to profile the bacterial, archaeal, and fungal communities across three root compartments for plants cultivated in three different soil types. Bars show the mean relative abundance (%) of each bacterial, fungal, or archaeal taxon within the endosphere, rhizosphere, or bulk soil of wheat cultivated in agricultural soil, Levington F2, compost or a 50:50 mix of the two (N=3 replicate plants per treatment). Colours indicate different microbial taxa (bacterial, fungal, or archaeal). For the archaeal community, N=2 replicate plants for the endosphere of plants grown in Levington F2 compost. Within stacked bars taxa are shown in reverse alphabetical order (left to right). The “Other” category includes all taxa with a median relative abundance of 0.05% or less. ASVs were assigned and are presented to the family level, where family-level taxonomic assignments were unavailable the next highest taxonomic assignment was presented.

**Table 4.1. Permanova results**

Permanova (permutational analysis of variance) comparing the bacterial, archaeal, or fungal community composition between conditions. The left panel shows comparisons between stem elongation growth phase field cultivated plants and laboratory cultivated four-week-old plants, and this showed no significant difference between communities in either set of plants. Right shows the effect of soil type on community composition for all three groups and compartments, and this showed in most cases a significant effect of soil type on community composition. N=3 for all comparisons, except for the archaeal endosphere community for Levington compost plants, where N=2.

Stem elongation compared to pot agricultural				Test for effect of soil type on community composition			
Community = Archaea							
Compartment	Permutations	R <sup>2</sup>	p-value	Compartment	Permutations	R <sup>2</sup>	p-value
Bulk Soil	999	0.12	0.7	<b>Bulk Soil</b>			
Rhizosphere	999	0.12	0.7	Archaea	999	0.94	0.003
Endosphere	999	0.34	0.2	Bacteria	999	0.87	0.001
<b>Community = Bacteria</b>				Fungi	999	0.81	0.004
Bulk Soil	999	0.3	0.1	<b>Rhizosphere</b>			
Rhizosphere	999	0.58	0.1	Archaea	999	0.97	0.004
Endosphere	999	0.53	0.1	Bacteria	999	0.83	0.001
<b>Community = Fungi</b>				Fungi	999	0.66	0.004
Bulk Soil	Fungi	0.53	0.1	<b>Endosphere</b>			
Rhizosphere	999	0.39	0.1	Archaea	999	0.87	0.004
Endosphere	999	0.42	0.1	Bacteria	999	0.6	0.001

## 4.2 Soil as a primary driver of root community diversity

To determine if the enrichment of specific microbial taxa and, in particular, the dominance of *Streptomyetaceae* and *Burkholderiaceae*, within the wheat root endosphere was driven by the soil community or by the host, *T. aestivum* var. Paragon was grown in the contrasting soil types (agricultural soil or compost), and a 50:50 mixture of the two. It was reasoned that if *Streptomyetaceae* and *Burkholderiaceae* were dominant only in the agricultural soil and the mixed soil, then

certain strains within the agricultural soil might be particularly effective at colonising the endosphere. However, if *Streptomycetaceae* and *Burkholderiaceae* were dominant in the endosphere across all three soil conditions, this would indicate that when present, this family is selectively recruited to the wheat root microbiome from diverse soil environments. The microbiome was compared between four-week-old (root growth phase) plants grown in Church Farm agricultural soil, Levington F2 compost, and a 50:50 (vol/vol) mix of the two soils under laboratory conditions. Church Farm soil and Levington F2 compost are starkly contrasting soil environments: the agricultural soil is mildly alkaline (pH 7.97), contained only 2.3% organic matter and was relatively low in inorganic nitrogen, magnesium, and potassium. Levington F2 compost is acidic (pH 4.98) and has a high organic matter content (91.1%) as well as higher levels of inorganic nitrogen, phosphorus, potassium, and magnesium (Chapter 2 Table 2.3).

**Table 4.2. SIMPER outputs**

Similarity percentages breakdown (SIMPER <sup>294</sup>) is a statistical package that scores the percent contribution of each microbial taxa to a community level change. This table shows the percent contribution (% contrib) of each family which significantly (see p-value) contributed to the shift in community composition from the bulk soil to rhizosphere, or bulk soil to endosphere across all soil types (N=3 per compartment).

Pot grown wheat, all soil types					
Comparison- Bulk soil-Endosphere			Bulk soil-Rhizosphere		
Taxa	Percent contribution	p-value	Taxa	Percent contribution	p-value
<i>Streptomycetaceae</i>	14.6	0.0099	<i>Burkholderiaceae</i>	8.9	0.4851
<i>Burkholderiaceae</i>	10.1	0.0099	<i>Rhizobiaceae</i>	5.4	0.0099
<i>Acidobacteria</i> Subgroup 6	3.6	0.0099	<i>Acidobacteria</i> Subgroup 6	4	0.0198
<i>Chitinophagaceae</i>	3.5	0.0792	<i>Bacillaceae</i>	3.4	0.0594
<i>Bacillaceae</i>	2.8	0.0099	<i>Micrococcaceae</i>	3.1	0.0297
<i>Xanthobacteraceae</i>	2.6	0.0099	<i>Chitinophagaceae</i>	2.6	0.9108
<i>Rhizobiaceae</i>	2.4	0.6733	<i>Xanthobacteraceae</i>	2.5	0.7227
Solirubrobacterales	2	0.0099	<i>Saccharimonadale</i>	2.4	0.3168
Saccharimonadales	2	0.0792	<i>Rubritaleaceae</i>	2.3	0.0099
<i>Methyloligellaceae</i>	1.6	0.0099	<i>Pseudomonadaceae</i>	2.2	0.0099

It is well documented that the soil microbial community is a major determinant of endosphere community composition, as endophytic microbes are acquired by plants horizontally from the soil <sup>52</sup>. The present study corroborates this observation as

PCoA showed clear clustering of communities by soil type, indicating that soil type was an important determinant of the root-associated community composition (Figure 4.1). For the bacterial and archaeal communities, permanova corroborated a significant effect of soil type on bacterial community composition for all compartments (Table 4.1). For the fungal community, permanova also showed significant effect of soil type on the bulk soil and rhizosphere communities (Table 4.1). For plants cultivated in Levington F2 compost, no data on the fungal community composition within the endosphere could be retrieved. Thus, no statistical comparison could be made. The bacterial communities were distinct between the bulk soil, rhizosphere, and endosphere. This indicated that, while the soil had a significant impact on the composition of the root associated communities, the plant also selects for specific microbial taxa in all the tested soils (Figure 4.1). PCoA showed a detectable rhizosphere effect (Figure 4.1) but, consistent with previous studies<sup>68,97</sup>, we observed a rhizosphere effect for *T. aestivum* var. Paragon that was subtle as there were only minor differences between the community composition of bulk soil and rhizosphere communities (Figure 4.2).

Regardless of soil type, *Streptomycetaceae* were amongst the bacterial taxa with the greatest average relative abundance in the endosphere (12.1% in compost, 16.9% in agricultural soil, and 39.3% in the 50:50 mixture), and were the most abundant endosphere taxon within the 50:50 soil mixture. The bacterial family showing the greatest average relative abundance within the endosphere of plants cultivated in agricultural soil was *Chitinophagaceae* (17%), in compost or the 50:50 mix however this family was a much lower proportion of the community (0.3% and 0.8% respectively). *Burkholderiaceae* showed the greatest average relative abundance within the endosphere of plants cultivated in compost (24.9%) and was amongst the top three endosphere taxa for agricultural soil or the 50:50 mix (14.3% and 15.9% respectively). *Burkholderiaceae* also showed the greatest average relative abundance within the rhizosphere within all three soils (agricultural soil 7.8%, compost 24.9% and 50:50 mix 8%). In Levington F2 compost the families *Rhizobiaceae* (13.2%) and *Soilrubrobacterales* (5.9%) showed the next-greatest average relative abundance. For agricultural soil however these taxa were *Rubritaleaceae* (7.3%) and *Micrococcaceae* (5.9%). In the 50:50 mixture of the two soils, *Rhizobiaceae*, *Soilrubrobacterales* and *Micrococcaceae* were amongst the taxa showing the greatest average relative abundance within the rhizosphere. *Xanthobacteraceae* (6.1%) and *Chitinophagaceae* (5.1%) were the second greatest proportion of the community. Conversely *Streptomycetaceae* were a low proportion

of the rhizosphere community (agricultural soil 2.1%, compost 2.1% and 50:50 mix 1.7%).

For the fungal community within plants cultivated in agricultural soil, the endosphere was colonised to the greatest extent by Ascomycota phylum (26.2%), followed by the Sordariomycetes order (25.2%) and the Pleosporales order (16.6%). For compost no data could be retrieved for the endosphere, though for the 50:50 mix *Metarhizum* showed the greatest average relative abundance within the endosphere (23.2%), followed by the orders Sordariomycetes (20.3%) and Agaricales (6.7%). In the rhizosphere Ascomycota showed the greatest average relative abundance (23.7%), followed by the family *Myxotrichaceae* (14.6%), and *Metarhizum* (14.6%). Similarly, *Metarhizum* showed the greatest average relative abundance within the rhizosphere of plants cultivated in agricultural soil (28.3%), followed by Sordariomycetes (22.8%), and Ascomycota (18.2%). For Levington F2 compost Agaricales dominated the rhizosphere community (44.9%), followed by Ascomycota (19%) and *Myxotrichaceae* (11.3%). The archaeal community was dominated by AOA; in agricultural soil *Nitrososphaeraceae* comprised 88.8% of the endosphere community whereas *Nitrosotaleaceae* were dominant within plants cultivated in Levington F2 compost (76.3%). In the 50:50 mixture *Nitrososphaeraceae* still dominated the archaeal endosphere community (78%), though *Nitrosotaleaceae* showed the second greatest average relative abundance (14.5%).

A SIMPER test revealed that, regardless of soil type, *Streptomycetaceae* (14.6% contribution,  $p < 0.01$ ) and *Burkholderiaceae* (10.1% contribution,  $p < 0.01$ ) were the main taxa driving the community shift from bulk soil to endosphere (Table 4.2). This is supported by the fact that *Streptomycetaceae* and *Burkholderiaceae* were major components of the endosphere bacterial communities under all conditions (Figure 4.2). Differential abundance analysis demonstrated a significant increase in the abundance of bacterial families *Burkholderiaceae*, *Chitinophageaceae*, *Pseudomonadaceae*, *Rhizobiaceae*, and *Streptomycetaceae* within the rhizosphere and/or endosphere across all soil types (Figure 4.3). Enrichment of these groups was correlated with the reduced abundance of some fungal taxa loosely associated with pathogenicity within the endosphere and rhizosphere (*Australiascaceae*<sup>331</sup>, *Glomerellaceae*<sup>332,333</sup> and *Hypocreales*<sup>334</sup>), and an increased abundance of one taxon loosely associated with beneficial mycorrhiza (*Leotiaceae*<sup>335-337</sup>) (Figure 4.3). For the fungal community however, this data did not include the endosphere compartment for Levington compost cultivated plants as no data was retrieved from

this compartment. Thus, this could also be responsible for these observed differential abundances.

qPCR experiments were performed to compare the abundance of archaea, bacteria, and fungi within the roots of plants cultivated in the agricultural soil or Levington F2 compost. No significant effect of soil type was observed for either fungal or bacterial abundance (ANOVA,  $p > 0.05$  for both) (Figure 4.). However, soil type had a significant effect on the abundance of archaea ( $p < 0.001$ ), and there were significantly greater numbers of archaea within the agricultural bulk soil and rhizosphere compartments when compared to Levington F2 compost (Tukey's HSD,  $p < 0.001$  for both), but there was no significant difference in the archaeal load detected within the endosphere (Tukey's HSD,  $p > 0.05$ ). The lower abundance of

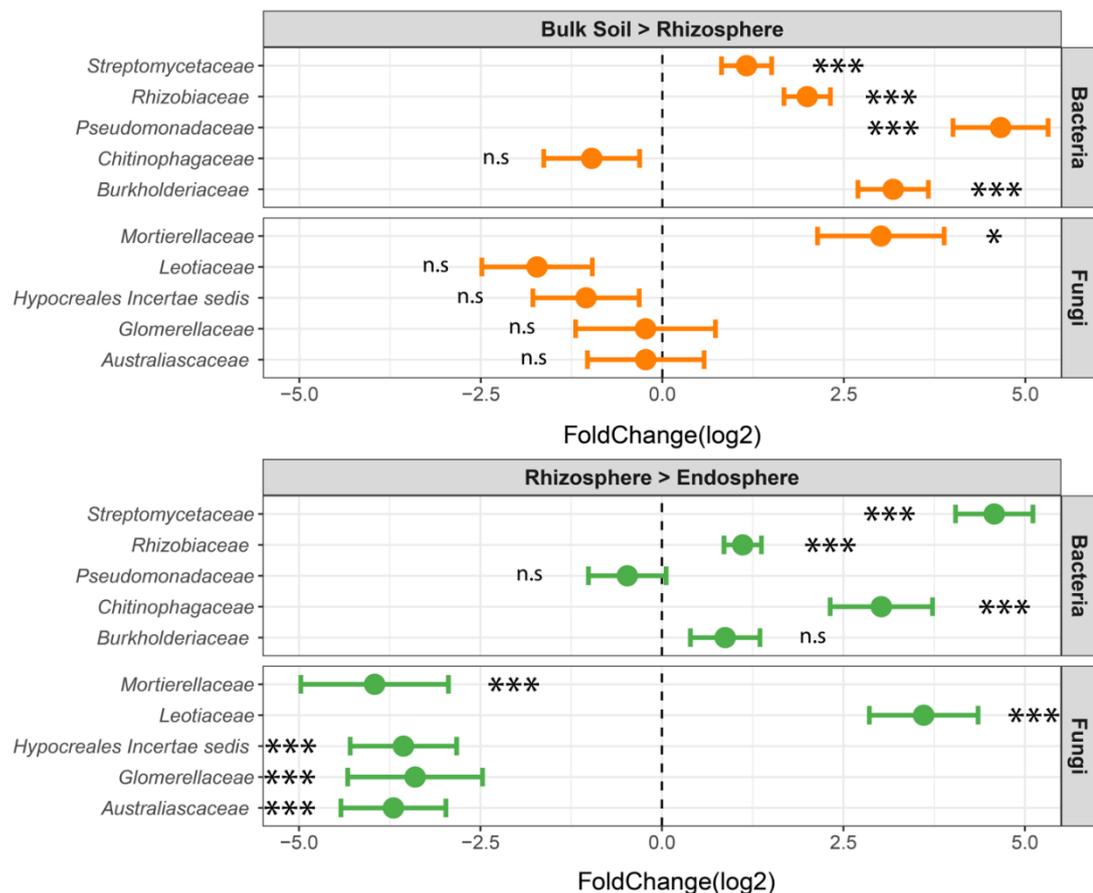


Figure 4.3 Differential abundance analysis performed to identify fungi or bacteria enriched within the rhizosphere or endosphere of wheat regardless of soil type. Dots show the log<sub>2</sub> fold change of different bacterial or fungal families and error bars show ± log fold change standard error. Results are from N=3 replicate plants per treatment. Shown are Bacterial and fungal taxa that were differentially abundant regardless of soil type for pot grown wheat. analysis was performed using DESeq2, \* indicates  $p < 0.05$ , \*\* indicates  $p < 0.01$ , \*\*\* indicates  $p < 0.001$ , and n.s. indicates  $p > 0.05$ , Bonferroni corrected. Data for all pot-grown plants were pooled and taxa which still showed significant fold change across compartments were included. For all complete statistical outputs see Supplementary Tables S.3 and S.7. For the fungal community no data was retrieved for the endosphere of Levington compost cultivated plants, so when comparing the rhizosphere and endosphere compartments, differential abundance is only within agricultural soil and 50:50 mix cultivated plants

archaea within Levington F2 compost is surprising given the higher nutrient levels in this soil, and particularly given the higher levels of ammonium (Table 2.3).

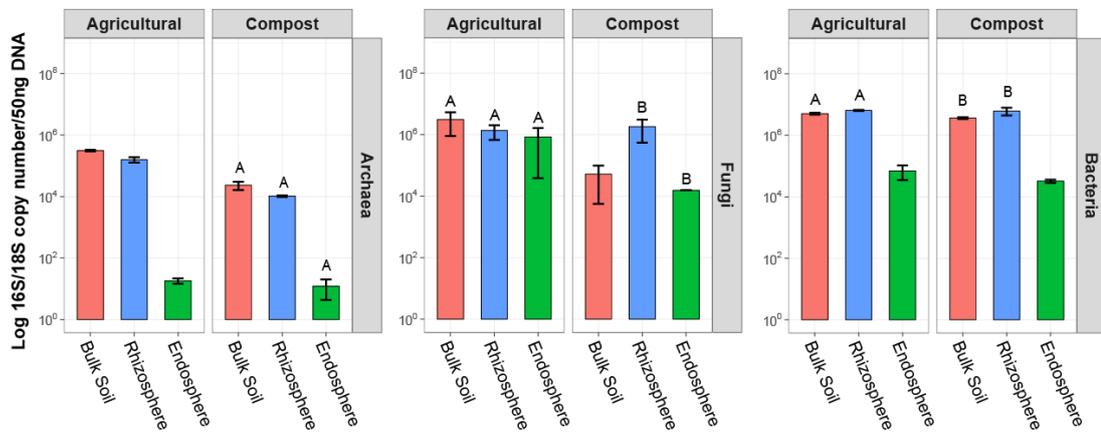


Figure 4.4 qPCR performed to quantify the absolute abundance of archaea, bacteria, or fungi across all three root compartments for plants cultivated in Levington F2 compost or agricultural soil. Bars show the mean log 16S or 18S rRNA gene copy per 50 ng of DNA within the endosphere, rhizosphere or bulk soil compartment of plants. Plants were grown in agricultural soil or compost, N=3 replicate plants per condition. Error bars show  $\pm$  standard error of the mean. For each group (archaea, bacteria or fungi) a general linear model (GLM) was used to test for a general effect of growth phase or compartment on 16S or 18S rRNA gene abundance. As a post-hoc a Tukey test was ran to identify significant differences between individual conditions. Comparisons were made between each compartment from either agricultural soil or Levington compost cultivated plants, and individual compartments were compared across soil conditions. For all three groups, samples labelled A or B showed no significant differences (Tukeys HSD,  $p > 0.05$ ), for all other comparisons a significant difference was found (Tukeys HSD,  $p < 0.05$  for all).

#### 4.3 Is *Triticum aestivum* var. Paragon an outlier among UK elite spring bread wheat?

Chapters 3.1 and 4.2 have demonstrated that some core microbiota, namely *Streptomyetaceae* and *Burkholderiaceae*, are consistently recruited to the root endosphere of Paragon var. *T. aestivum*. There is some evidence however that Paragon is an outlier amongst UK elite spring bread wheat varieties, and thus other varieties may not be enriched for these taxa within the endosphere<sup>97</sup>. The Paragon variety of UK elite spring bread wheat, along with the variety Cadenza, is a cross between parent varieties Tonic and Axona<sup>338</sup>. To explore the consistency of the microbial community of wheat across genotypes, five varieties of UK elite spring bread wheat were cultivated in agricultural soil sampled from the Church Farm site. These varieties included Cadenza and Paragon, the two parental varieties Axona and Tonic, and a distantly related UK wheat variety, Soissons<sup>338</sup>. The relationship between these wheat varieties is demonstrated in Figure 4.6, which shows the relatedness of all UK wheat varieties. This shows that Cadenza, Paragon, Axona, and Tonic are all closely related, whilst Soissons is derived from a distantly related lineage.

The same three taxa showed the greatest average relative abundance within the endosphere of all five varieties: *Streptomycetaceae* (Axona 24.1%, Cadenza 25.2%, Paragon 15.4%, Soissons 17.2%, Tonic 21.6%), *Burkholderiaceae* (Axona 16.6%, Cadenza 17.2%, Paragon 12.8%, Soissons 20.7%, Tonic 20.3%), and *Chitinophagaceae* (Axona 7%, Cadenza 8.1%, Paragon 7.1%, Soissons 6.1%, Tonic 7.5%). Contrary to previous experiments presented in this work, *Bacillaceae* were also among the taxa showing the greatest relative abundance within the rhizosphere (Axona 8.8%, Cadenza 9.4%, Paragon 10.4%, Soissons 8%, Tonic 7.8%), likely due to variation from soil sampling occurring at different times of the year. This was followed by *Chitinophagaceae* (Axona 8.3%, Cadenza 4.4%, Paragon 2.4%, Soissons 3.4%, Tonic 3.4%), and *Burkholderiaceae* (Axona 7.6%, Cadenza 6.1%, Paragon 6.5%, Soissons 8.9%, Tonic 8%). For Cadenza and Soissons, *Xanthobacteraceae* showed the third-greatest average relative abundance within the rhizosphere (Cadenza 5.3%, Soissons 5.5%), the equivalent for Paragon was *Rhizobiaceae* (4.9%) and for Tonic was *Methyloligellaceae* (5%) (Figure 4.5). Interestingly *Bacillaceae* were also a larger proportion of the bulk soil community than was observed in previous experiments, despite the soil for all experiments being sampled from the same site. *Bacillaceae* were on average 11.1% of the bulk soil community across the five varieties, compared to an average of 6% within bulk soil across the previous three experiments (Chapter 3.1, 3.2 and 4.2). This difference is likely attributable to different seasons used for soil collection across experiments and may explain the higher proportion of *Bacillaceae* detected within the rhizosphere of the different wheat varieties in the present experiment.

PCoA showed that regardless of variety, there was a clear distinction between the bulk soil, rhizosphere, and endosphere communities (Figure 4.7), demonstrating that all varieties influence the microbial community within the root. This was corroborated by permanova which showed a significant effect of compartment on community composition ( $R^2 = 0.61$ ,  $p = 0.0001$ ), though betadisper was significant ( $p = 0.0001$ ) which indicated that microbial community dispersion was not equal across the root compartments. PCoA suggested that this may be explained by inconsistent rhizosphere and endosphere replicates for Axona and Cadenza (Figure 4.7 A, All Samples). Permanova indicated a weak yet significant effect of variety on community composition ( $R^2 = 0.07$ ,  $p < 0.05$ , betadisper n.s). PCoA however did not show strong variety-based clustering of communities, except for within the endosphere of Axona and Cadenza, which weakly clustered separately to the other three varieties (Figure 4.7 A, Bulk Soil, Rhizosphere, Endosphere). Pairwise

permanova found no significant difference between the endosphere or rhizosphere communities associated with any two varieties of wheat (Table 4.3). Overall, this analysis indicated that there may be minor differences between the microbial communities associated with these five varieties of *T. aestivum*, but that, when cultivated within the same soil conditions, the same major taxa colonise the root community regardless of variety. Analysis of composition with bias correction (ANCOM-BC<sup>297</sup>) indicated that at the family level five taxa above the abundance threshold (0.05% median relative abundance) were significantly differentially abundant within the endosphere across varieties (Table 4.4). In the rhizosphere ANCOM-BC indicated that just two taxa were significantly differentially abundant across varieties (Table 4.5).

Above the abundance threshold (0.05% median relative abundance) ANCOM-BC indicated that 16 taxa were enriched within the endosphere or rhizosphere across all five wheat varieties (Table 4.6). This included all of the core enriched taxa identified in Chapter 4.2, *Burkholderiaceae*, *Chitinophageaceae*, *Pseudomonadaceae*, *Rhizobiaceae*, and *Streptomycetaceae*. This indicates that not only can wheat recruit these families to the root community from contrasting soil environments, but that a range of different UK wheat varieties can recruit these taxa. This might indicate a broadly conserved interaction between these bacterial families and UK wheat varieties, particularly given that the core taxa were found not only in association with the closely related varieties (Axona, Tonic, Paragon, and Cadenza) but also the more distantly related variety<sup>338</sup>, Soissons.



Figure 4.5. Metabarcoding performed to profile the bacterial community across three root compartments for five different varieties of UK elite spring bread wheat. Bars show the mean relative abundance (%) of each bacterial taxon within the endosphere, rhizosphere or bulk soil of wheat cultivated in agricultural soil (N=3 replicate plants per variety). Five varieties are shown, Axona, Cadenza, Paragon, Soissons, and Tonic. Colours indicate different microbial taxa. Within stacked bars taxa are shown in reverse alphabetical order (left to right). The “Other” category contains all taxa falling below the median abundance threshold of 0.05%. The faint lines within the bars indicate different ASVs within each taxon.

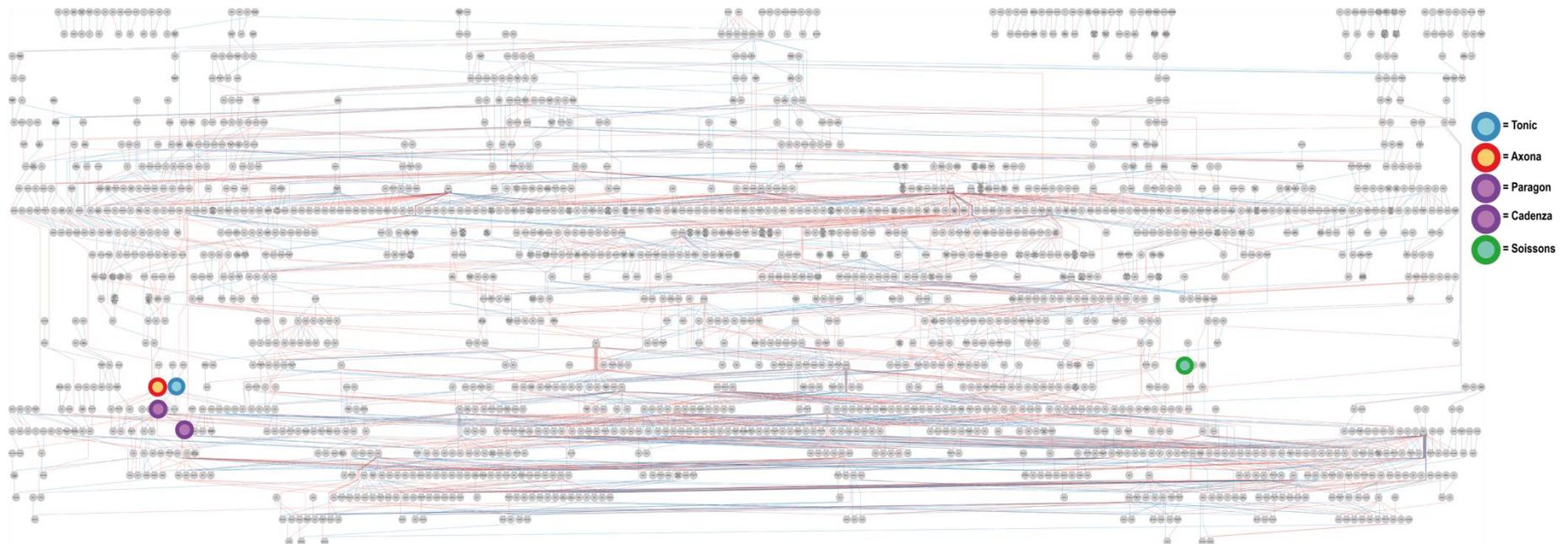


Figure 4.6. The UK wheat varieties pedigree <sup>338</sup>, showing the relatedness of all UK wheat varieties. This was produced using the software Helium to demonstrate the relatedness of five different varieties of wheat used in this work. Lines show parentage of different wheat varieties, blue denoting male and red denoting female parentage. Circles show the location of varieties Tonic (blue), Axona (red), Paragon and Cadanza (purple), and Soissons (green).

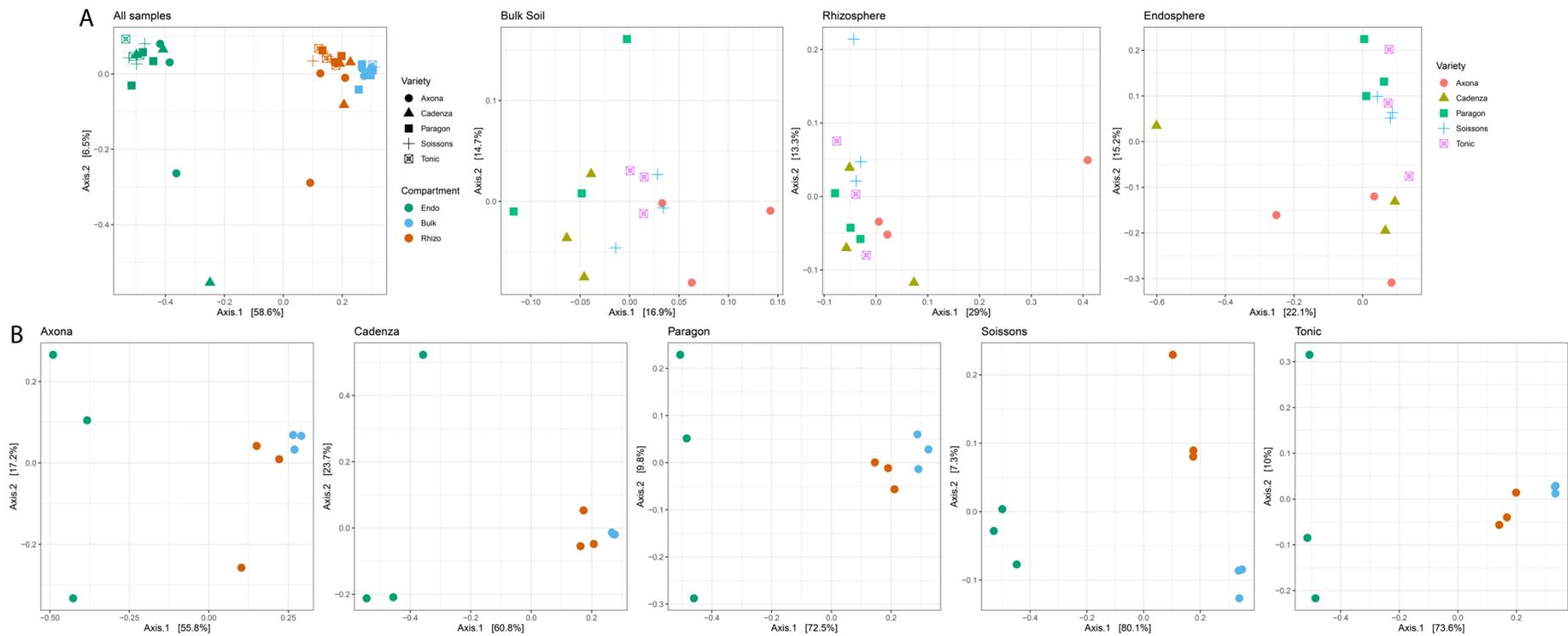


Figure 4.7 Principal Coordinates Analysis (PCoA) performed on Bray Curtis dissimilarities to assess the similarity of bacterial communities associated with each root compartment for five varieties of UK elite spring bread wheat. **A** PCoA comparing the bacterial community in the bulk soil, rhizosphere, and endosphere associated with different wheat varieties. In the plot titled **All Samples** colours indicate root compartment; green = endosphere, orange = rhizosphere, and blue = bulk soil. Shapes indicate different varieties of wheat; circle = Axona, triangle = Cadenza, square = Paragon, cross = Soissons, cross-box = Tonic. In the plots titled **Bulk Soil**, **Rhizosphere** and **Endosphere**, colours also indicate variety; Red = Axona, Olive = Cadenza, Green = Paragon, Blue = Soissons, pink = Tonic. **B** PCoA comparing the bacterial community across root compartments for each individual variety. Colours indicate root compartment; green = endosphere, orange = rhizosphere, and blue = bulk soil. For all varieties N=3 replicate plants per condition.

**Table 4.3. Pairwise permanova outputs**

Permanova (permutational analysis of variance) comparing the bacterial endosphere or rhizosphere communities between five different varieties of UK elite spring bread wheat (N=3).

Comparison	Endosphere		Rhizosphere	
	R <sup>2</sup>	P-value	R <sup>2</sup>	P-value
Axona - Cadenza	0.17261	0.6	0.24856	0.2
Axona - Paragon	0.27914	0.1	0.29431	0.1
Axona - Tonic	0.25057	0.1	0.27115	0.1
Axona - Soissons	0.32358	0.1	0.29527	0.1
Cadenza - Paragon	0.25777	0.1	0.2662	0.2
Cadenza - Tonic	0.23307	0.2	0.27504	0.1
Cadenza - Soissons	0.29331	0.1	0.34197	0.1
Paragon - Soissons	0.31908	0.1	0.33908	0.1
Paragon - Tonic	0.21191	0.4	0.21605	0.3
Tonic - Soissons	0.28633	0.1	0.29447	0.1

**Table 4.4. ANCOM-BC results, endosphere**

ANCOM-BC (analysis of composition with bias correction) performed to identify bacterial taxa which were significantly differentially abundant within the endosphere of any of the five varieties of UK elite spring bread wheat assessed, when compared to Paragon (N=3).

Family	Coefficient	Test Statistic	P-adjust	Standard Error	Comparison
<i>Bacillaceae</i>	1.124	5.420577761	0.0000117	0.207380339	Paragon - Axona
	0.868	3.815014768	0.02587321	0.227590648	Paragon - Cadenza
<i>Devosiaceae</i>	-1.101	-4.48017806	0.001439409	0.2456517	Paragon - Axona
	-0.918	-3.473968874	0.096410257	0.264168206	Paragon - Cadenza
	-0.879	-6.237227707	0.0000000864	0.141013036	Paragon - Tonic
<i>Hyphomicrobiaceae</i>	-1.120	-15.57231046	2.21x10 <sup>-51</sup>	0.071907272	Paragon - Tonic
<i>Microscillaceae</i>	-0.966	-3.782458884	0.029349196	0.255384047	Paragon - Cadenza
<i>Nocardiaceae</i>	0.773	4.673306274	0.000577959	0.165421067	Paragon - Soissons
	0.633	4.746602627	0.000395106	0.133309747	Paragon - Tonic

**Table 4.5. ANCOM-BC results, rhizosphere**

ANCOM-BC (analysis of composition with bias correction) performed to identify bacterial taxa which were significantly differentially abundant within the rhizosphere of any of the five varieties of UK elite spring bread wheat assessed, when compared to Paragon (N=3).

Family	Coefficient	Test Statistic	P-adjust	Standard Error	Comparison
<i>Fibrobacteraceae</i>	-0.901	-7.8004800	0.000000000000149	0.11553827	Paragon - Axona
<i>Spingobacteriaceae</i>	-0.511	-4.0586308	0.011747984	0.474687671	Paragon - Tonic

**Table 4.6. ANCOM-BC results, cross-compartment**

ANCOM-BC performed to identify bacterial taxa which were significantly differentially abundant within the endosphere compared to the rhizosphere, or rhizosphere compared to bulk across all five of the UK wheat varieties assessed (N=3).

Family	Coefficient	Test Statistic	P-adjust	Standard Error	Comparison
<i>Burkholderiaceae</i>	2.5310333	12.74898007	2.61x10 <sup>-35</sup>	0.198528297	Rz - En
	1.9682827	14.41169717	1.07x10 <sup>-44</sup>	0.136575357	BS - Rz
<i>Caulobacteraceae</i>	2.2311904	9.653333389	1.11x10 <sup>-19</sup>	0.231131602	Rz - En
	1.3262715	9.229454187	6.61x10 <sup>-18</sup>	0.14369988	BS - Rz
<i>Chitinophagaceae</i>	2.2804867	7.767149779	1.78x10 <sup>-12</sup>	0.293606632	Rz - En
<i>Devosiaceae</i>	1.7529109	8.404213896	9.73x10 <sup>-15</sup>	0.208575239	Rz - En
	1.2547316	9.901944854	9.96x10 <sup>-21</sup>	0.126715671	BS - Rz
<i>Enterobacteriaceae</i>	3.7938978	5.318339799	0.0000194776	0.713361304	Rz - En
	3.6316204	5.111804036	0.0000717968	0.710438109	BS - Rz
<i>Fibrobacteraceae</i>	2.2718143	8.610880229	1.65x10 <sup>-15</sup>	0.263830672	Rz - En
	0.8386932	6.674267741	0.0000000058	0.125660712	BS - Rz
<i>Gaiellaceae</i>	-1.6873931	-4.130508117	0.0056954509	0.4085195	Rz - En
	-0.2490370	-4.510496868	0.001409936	0.055212754	BS - Rz
<i>Gemmatimonadaceae</i>	-2.7385881	-7.037956443	0.0000000004	0.389116942	Rz - En
	-0.2454850	-4.229949995	0.0050488545	0.058034961	BS - Rz
<i>Methyloigellaceae</i>	-0.6602540	-5.614129974	0.000003793	0.117605757	Rz - En
	-0.2968459	-4.072791697	0.0099409349	0.072885116	BS - Rz
<i>Micromonosporaceae</i>	3.5701989	13.35351582	2.76x10 <sup>-38</sup>	0.267360221	Rz - En
<i>Microscillaceae</i>	2.0108652	9.779447954	2.76x10 <sup>-20</sup>	0.20562154	Rz - En
	0.5109871	6.092354268	0.000000257	0.083873507	BS - Rz
<i>Mycobacteriaceae</i>	-1.6309835	-4.05201356	0.0076675869	0.402511867	Rz - En
<i>Nitrospiraceae</i>	-2.0488011	-8.727598169	5.98x10 <sup>-16</sup>	0.2347497	Rz - En
<i>Nocardiaceae</i>	0.0516607	4.391072059	0.002447611	0.129069537	BS - Rz
<i>Paenibacillaceae</i>	1.7226820	11.50840834	2.88x10 <sup>-28</sup>	0.149688987	Rz - En
<i>Pseudomonadaceae</i>	2.4568329	6.413784703	0.0000000298	0.383055096	Rz - En
	1.4684063	7.802616078	1.46x10 <sup>-12</sup>	0.188194097	BS - Rz
<i>Pseudonocardiaceae</i>	3.2815983	10.92308968	2.15x10 <sup>-25</sup>	0.300427667	Rz - En
<i>Pyrimonadaceae</i>	-2.3132688	-6.120725421	0.0000001872	0.377940305	Rz - En
<i>Rhizobiaceae</i>	1.4502132	6.015703339	0.0000003528	0.241071262	Rz - En
	0.7131365	-4.870842736	0.0002466936	0.146409262	BS - Rz
<i>Rhizobiales Incertae Sedis</i>	-2.0986457	-6.098653956	0.0000002139	0.344116207	Rz - En
<i>Solibacteraceae</i> (Subgroup 3)	-2.5319113	-5.576702738	0.0000046818	0.454015828	Rz - En
<i>Sphingobacteriaceae</i>	1.6617956	4.191056002	0.0044703117	0.396509992	Rz - En
	1.0675991	-6.248698687	0.0000000964	0.170851422	BS - Rz
<i>Streptomycetaceae</i>	4.0247648	18.22111509	8.67x10 <sup>-72</sup>	0.220884658	Rz - En
<i>Xanthomonadaceae</i>	1.7348821	5.418165838	0.0000113803	0.320197308	Rz - En

BS = Bulk Soil, Rz = Rhizosphere, En = Endosphere

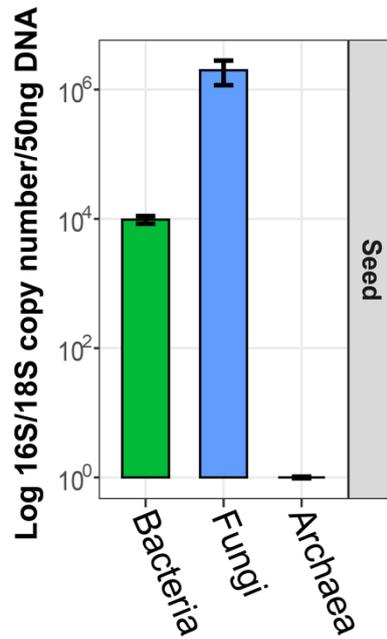


Figure 4.8 qPCR experiment performed to assess the absolute abundance of bacteria, fungi, or archaea within the interior of surface sterilized *T. aestivum* var. Paragon seeds. Bars show the mean log 16S or 18S rRNA gene copy per 50 ng of DNA, N=3 replicate DNA extracts per treatment (each extraction was performed on material pooled from five wheat seeds). Error bars show  $\pm$  standard error of the mean.

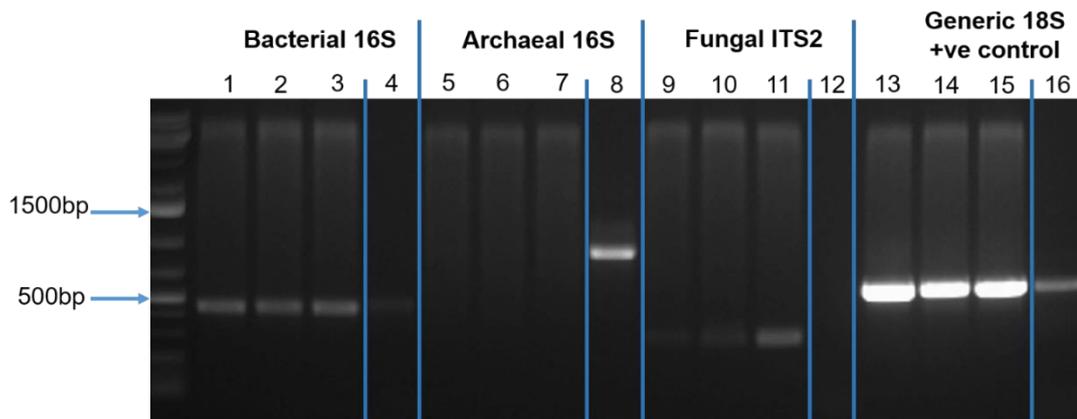


Figure 4.9 Agarose gel electrophoresis to visualise PCR amplification of a bacterial, archaeal, or fungal barcoding region from DNA extracted from the interior of wheat seeds. Lanes 1-4 show bacterial 16S rRNA gene amplification using primers PRK341F/MPRK806R (expected size 430 bp), lanes 5-8 show archaeal 16S rRNA gene amplification using primers A108F/A1000R (expected size, 1000 bp), lanes 9-12 show fungal ITS2 region amplification using primers ITS7F/ITS4R\_2 (expected size 500bp), and lanes 13-16 show generic 18S rRNA gene amplification using primers 18S1A/18S564R as a positive control (expected size 560 bp). Lanes 4, 8, 12, and 16 show amplification for each primer set from soil DNA as a positive control, this showed successful amplification for all primer pairs except for the fungal ITS2 region.

#### 4.4 The endophytic seed microbiome

Though it is generally accepted that root associated microbiomes are acquired horizontally from the soil <sup>52</sup>, there is some debate surrounding the extent to which the seed endophytic microbiome can influence microbial assemblages in the mature plant <sup>339</sup>. To investigate for *T. aestivum* var. Paragon, 16S rRNA gene and ITS2 gene sequencing was performed from the interior of surface sterilised wheat seeds, and qPCR was used to approximate the absolute abundance of these groups. qPCR showed that, while there were significant quantities of bacterial 16S rRNA gene or fungal 18S rRNA gene copies within the seeds, very few archaeal 16S rRNA gene copies could be detected (Figure 4.8). Indeed, when archaea-specific 16S rRNA gene primers were used on seed-interior DNA extracts, no visible amplification could be achieved (Figure 4.9). This strongly implies that there are no detectable archaea present inside of the seeds. For the other two groups however, significant quantities were detected within the seeds. While orders of magnitude lower than the quantities detected in the soil, the quantity of bacteria within the seeds was approximately  $9.7 \times 10^3$  copies / 50 ng DNA. For fungi, there were orders of magnitude more 18S rRNA gene copies within the seeds compared to bacteria, with  $2 \times 10^6$  copies / 50 ng of seed endosphere DNA (Figure 4.8).

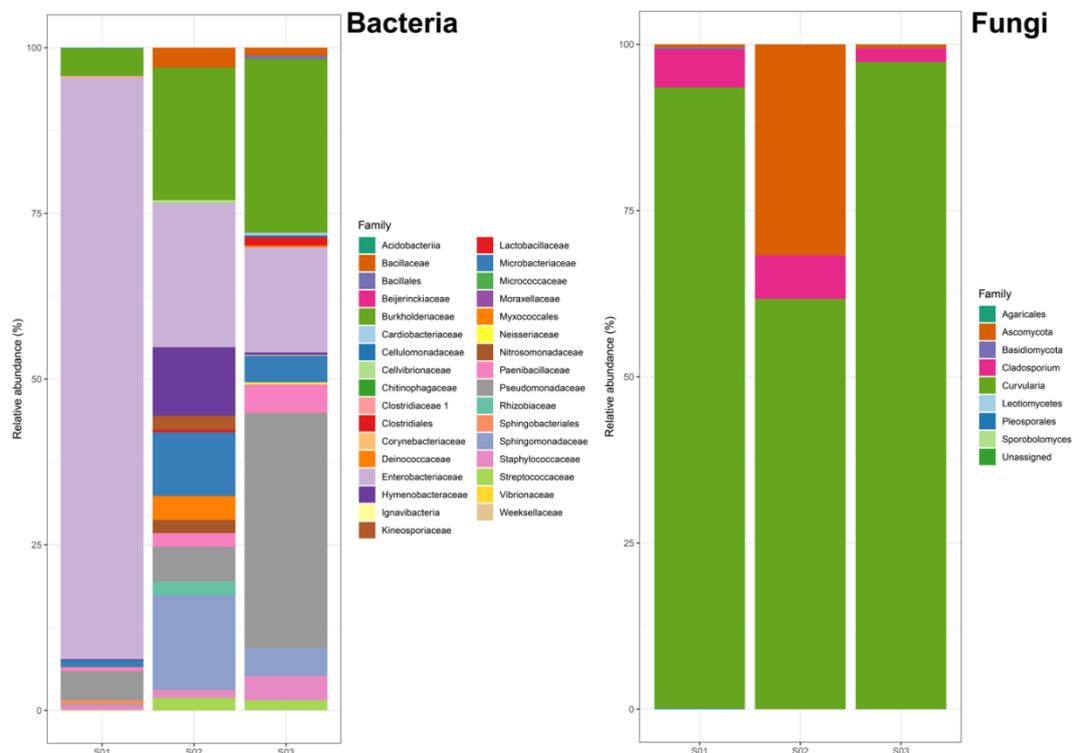


Figure 4.10. Metabarcoding performed to profile the bacterial and fungal community within wheat seeds. Bars show the relative abundance (%) of each bacterial (left) or fungal (right) taxon within the interior of surface sterilised seeds. Each bar is one replicate and DNA was extracted from five seeds per replicate. Colours indicate different microbial taxa, within stacked bars taxa are shown in alphabetical order (top to bottom).

Supplementary Figure S.3 shows that there was  $1.1 \times 10^4$  bacterial 16S rRNA gene copies / 50 ng DNA within the endosphere of wheat seedlings cultivated under sterile conditions. Plants cultivated under sterile conditions would usually be assumed to harbour no microbes within the roots or are often referred to as gnotobiotic. This test however appears to show that there is a detectable quantity of bacteria within the roots. In a sterile system the only possible source for these microbes is the seed endosphere, demonstrating that it is possible for bacteria to colonise the roots from the seeds. The bacterial load within aseptically cultivated seedling roots was four orders of magnitude fewer than that of four-week-old plants cultivated using potted agricultural soil (Figure 4.4). Whilst seedlings were used for this experiment (~1 week old), this might indicate that a large proportion of the bacterial community is missing as the plants were unable to recruit bacteria from the soil.

Sequencing of the bacterial 16S rRNA gene revealed a diverse range of bacteria within the seed interior, major families present included *Burkholderiaceae* (16.8%), *Enterobacteriaceae* (41.8%), *Hymenobacteraceae* (3.6%), *Microbacteriaceae* (4.9%), *Pseudomonadaceae* (15.1%), and *Sphingomonadaceae* (15.1%). Many of these taxa are core rhizosphere or endosphere bacteria, such as *Burkholderiaceae* and *Pseudomonadaceae*, and other core microbes present include *Chitinophagaceae* (0.02%) and *Rhizobiaceae* (0.7%) (Figure 4.10). The *Streptomyetaceae* family were not detected in any of the three seed samples and were the only core root associated taxon not to be detected within the seeds. The community varied across replicates; for example, while on average *Enterobacteriaceae* showed the greatest relative abundance, this varied from 87.7% to 15.9% of the community, and in one replicate *Pseudomonadaceae* showed the greatest relative abundance (35.5%). Overall, as demonstrated in Figure 4.10, consistency for the bacterial community was poor across replicates, indicating that the mechanism by which bacterial seed endosphere community is assembled may be stochastic, or sensitive to the variable environmental factors in the field from which these seeds were harvested.

For the fungal community only a small proportion the reads could not be classified (0.04%), and across the replicates the community was dominated by the genus *Curvularia* (84.1%). Other major groups included the phylum Ascomycota (10.9%) and the genus *Cladosporium* (4.8%). Interestingly the Leotiomycetes class was also detected within the seeds (0.001%), this contains the only core endosphere enriched fungal family, *Leotiaceae* (Figure 4.3).

#### 4.5 Long read amplicon sequencing for genus to species level identification of core endosphere bacteria

After five core root associated microbial taxa were identified, long read PacBio SMRT closed circular sequencing (CSS) was used to sequence the full length bacterial 16S rRNA gene, to identify these taxa from the genus to species level, and to understand variation within these groups. For the bulk soil, rhizosphere, and seed compartments very few sequences were recovered after quality filtering stages (118, 50, and 0 respectively). As a result, despite the recovery of 12,638 good quality full length 16S rRNA gene sequences from the endosphere, cross compartment comparisons for whole community composition were not deemed viable for this dataset.

For the *Burkholderiaceae* nine unique sequences were recovered from the endosphere, all of which were classified as members of the *Oxolabacteriaceae* and *Comamonadaceae* families now classified by the GTDB within the *Burkholderiaceae*<sup>304</sup>. Six of these sequences of the genus *Duganella* and clustered closely with example *Duganella* sequences (Figure 4.11). These six sequences collectively accounted for 1905 reads, and all six showed >98% sequence similarity with the example database sequence from *Duganella phyllosphaerae*. One *Duganella* 16S rRNA gene sequence was also recovered from the rhizosphere, and this showed 100% sequence identity with the second most numerous endosphere *Duganella* sequence (Supplementary Table S.9) showing that the same *Duganella* strain was detected within both the endosphere and rhizosphere. The second most numerous *Burkholderiaceae* taxon was *Massilia*, represented by 945 copies of a single unique 16S rRNA gene sequence within the endosphere, which showed >99% sequence identity with the example *Massilia* sequence. Two unique *Massilia* sequences were also recovered from the rhizosphere and showed >99% sequence identity with the endosphere *Massilia* sequence, indicating that the rhizosphere and endosphere strains are very closely related. When compared to the 16S rRNA gene sequence from the isolated *Massilia* strain (RNA126, see Chapter 6 for more detail) these sequences all had approximately 97% sequence similarity, strongly indicating these sequences originate from the same genus<sup>132</sup>. The third most numerous genus was *Rhodofera*, which showed 98% sequence similarity to the example database *Rhodofera* sequence *Rhodofera sediminis*, strongly indicating this recovered sequence is from the *Rhodofera* genus. Two other genera were detected within the endosphere, represented by a single unique sequence and with relatively low sequence counts, *Polaromonas* (28 sequences) and *Herbaspirillum* (10 sequences).

Within the rhizosphere two additional *Burkholderiaceae* family taxa were detected, *Rhodobacter* and *Variovorax*. The *Variovorax* sequences showed 98-99% sequence similarity with the four cultured representatives of this genus (see Chapter 6), indicating these cultured strains are closely related to the strains present within the rhizosphere. Overall, whilst a diverse community of *Burkholderiaceae* were present within the roots, the *Massilia* and *Duganella* genera accounted for the largest proportion of these reads, indicating these genera are both prevalent within the wheat root and are responsible for the enrichment of the *Burkholderiaceae* family within the endosphere.

For the *Streptomycetaceae* family six unique 16S rRNA gene sequences were recovered from the endosphere, accounting for 1326 reads. These clustered into three distinct clades (Figure 4.12). One of these clades contained just one unique 16S rRNA sequence of which there were 68 copies; within this group the recovered sequence shared 100% sequence identity with the database representative of the *Streptomyces canus* species, and also shared 100% sequence identity with four endosphere *Streptomyces* isolates (CRwSp2b, SRwSp1, PRwSp2, and PES2) (Supplementary Table S.10). This strongly indicated that these isolates and the recovered sequence belong to the species *Streptomyces canus* and are closely related to the database representative. One other clade contained just one of the recovered sequences, annotated by the SILVA database<sup>290</sup> as *Streptomyces scabiei*. For this unique 16S rRNA gene sequence just 25 copies recovered from the endosphere, and it did not align with 100% identity to any of the other sequences, including to the representative sequence for *S. scabiei* 87.22 acquired from the NCBI database with which it shared 99% sequence identity. This indicates that this sequence could belong to the *S. scabiei* species. The majority of the *Streptomycetaceae* reads (1233 of the sequences) belonged to four unique 16S rRNA gene sequences which were all grouped within the same clade, identified as *Streptomyces turgidiscabies* (Figure 4.12). All five of these sequences shared 99% sequence identity with the database representative for *S. turgidiscabies*, strongly indicating that they belong to this species. The most numerous of these *S. turgidiscabies* sequences that were recovered from the endosphere, representing 684 of the *Streptomycetaceae* sequences, showed 100% sequence identity with two of the *Streptomyces* endosphere isolates (CESp2 and CESp3), indicating that these isolates belong to this species and are possibly the same strain as the most numerous streptomycete within the endosphere.

For the *Pseudomonadaceae* just two sequences were recovered from the PacBio sequencing data, one from the rhizosphere, and one from the endosphere for which there was 191 sequences. Both of these sequences shared 99% sequence identity with the example sequence from *Pseudomonas poae* (Supplementary Table S.11), with which they also clustered (Figure 4.13), indicating the most prevalent *Pseudomonas* strains within the endosphere belonged to the *P. poae* species. All four *Pseudomonas* endosphere isolates however clustered with *Pseudomonas brassicacaerum* (Figure 4.13), with which they shared 99-100% sequence identity, whilst sharing 97-98% identity with *P. poae*. This indicates that these isolates are more closely related to the representative for *P. brassicacaerum* than to the *P. poae* strain that was detected within the endosphere.

For the *Rhizobiaceae* four unique taxa were recovered from the endosphere, representing a total of 231 of the endosphere derived sequences. The most numerous of these unique sequences were annotated as *Rhizobiaceae*, and clustered disparately (Figure 4.14). In a BLAST search using the NCBI database, the most abundant of these (93 reads) aligned with 100% similarity to *Agrobacterium rubi*. The second most abundant (85 reads) shared 100% sequence similarity with *Rhizobium giardinii*, and the least abundant aligned with >99% sequence identity to a number of *Rhizobium* species. Overall, this indicates the presence of *Agrobacterium* and *Rhizobium* genera bacteria within the endosphere. Seven *Rhizobiaceae* reads were attributed to the species *Neorhizobium galegae*; this sequence clustered with a representative of this species from the NCBI database, with which it shared 98% sequence identity. A BLAST search using the NCBI database<sup>292</sup> revealed a 99.71% sequence identity with a different *N. galegae* strain, strongly indicating these reads are derived from bacteria of this species. Thirteen reads were attributed to the *Phyllobacterium* genus, again this sequence clustered with the *Phyllobacterium* database representative, with which it shared >97% sequence identity (Supplementary Table S.12), indicating that *Phyllobacterium* genus bacteria were present within the root. A BLAST search using this sequence did not provide species level identification, implying that this specific *Phyllobacterium* species may remain undiscovered. Within the *Rhizobiaeae* a range of species were detected, none of which were overwhelmingly abundant in comparison to the others. This shows that a diverse community of genera within this family resides within the root, and that strain to species level changes in the abundance of bacteria within this group may be masked by the family level

identifications provided by the previously described short read sequencing experiments.

For the *Chitinophagaceae* four unique 16S rRNA gene sequences were recovered. Two of these (for which there were 148 reads in total) clustered with the database representative for the *Niastella* genus (Figure 4.15). These sequences however showed only 94-95% sequence identity with the *Niastella* database representative (Supplementary Table S.13). A BLAST search using the NCBI database did not yield any hits of a higher sequence identity than this. Nonetheless, we can be confident these two 16S rRNA sequences are derived from *Niastealla* genus bacteria<sup>132</sup>. One 16S rRNA sequence clustered with the database representative for the *Chitinophaga*, with which it showed >97% sequence identity, indicating that these reads are from a *Chitinophaga* genus bacterium. A BLAST search using the NCBI database revealed that this sequence shares >99% identity with the 16S rRNA gene for *Chitinophaga oryzae* ZBGKL4, so a strain from this species could be present within wheat roots. One additional sequence was recovered from the rhizosphere, and this showed 95% sequence identity with *Terrimonas ferruginea*, indicating that this genus is present within the rhizosphere. A BLAST search using the NCBI database did not reveal any more information about the potential identity of this sequence.

# Burkholderiaceae

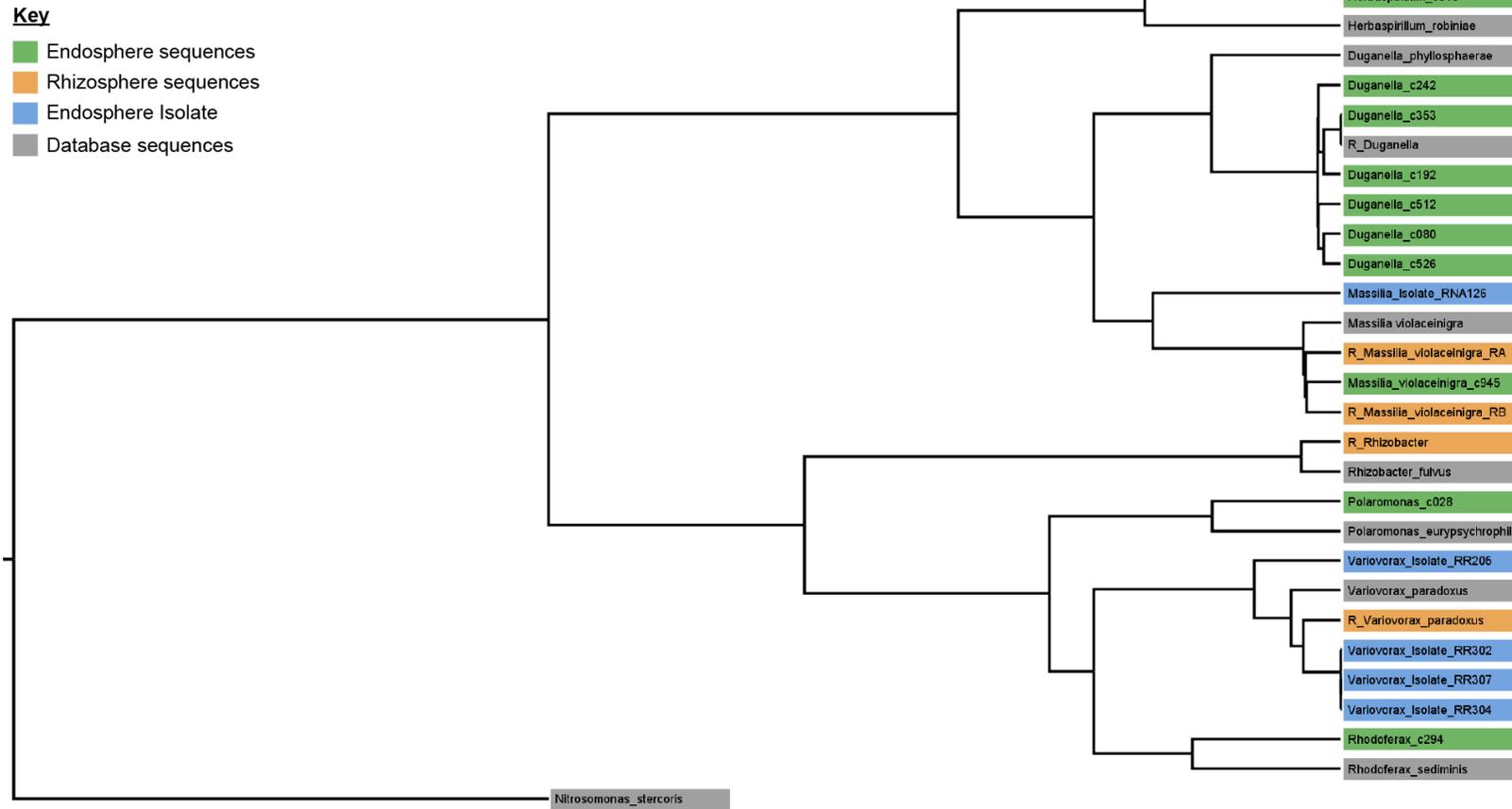


Figure 4.11. Phylogenetic tree showing the relatedness of *Burkholderiaceae* family 16S rRNA gene sequences recovered from the wheat root endosphere (green) or rhizosphere (orange), compared to *Burkholderiaceae* root isolates of the same genera (blue) (see chapter 6 for more details on isolates), and to example 16S rRNA gene sequences acquired from the NCBI database (grey). The tree was rooted using the 16S rRNA gene sequence of a related taxa from a different family within the Burkholderiales order, *Nitrosomonas stercoris*.

# Streptomycetaceae

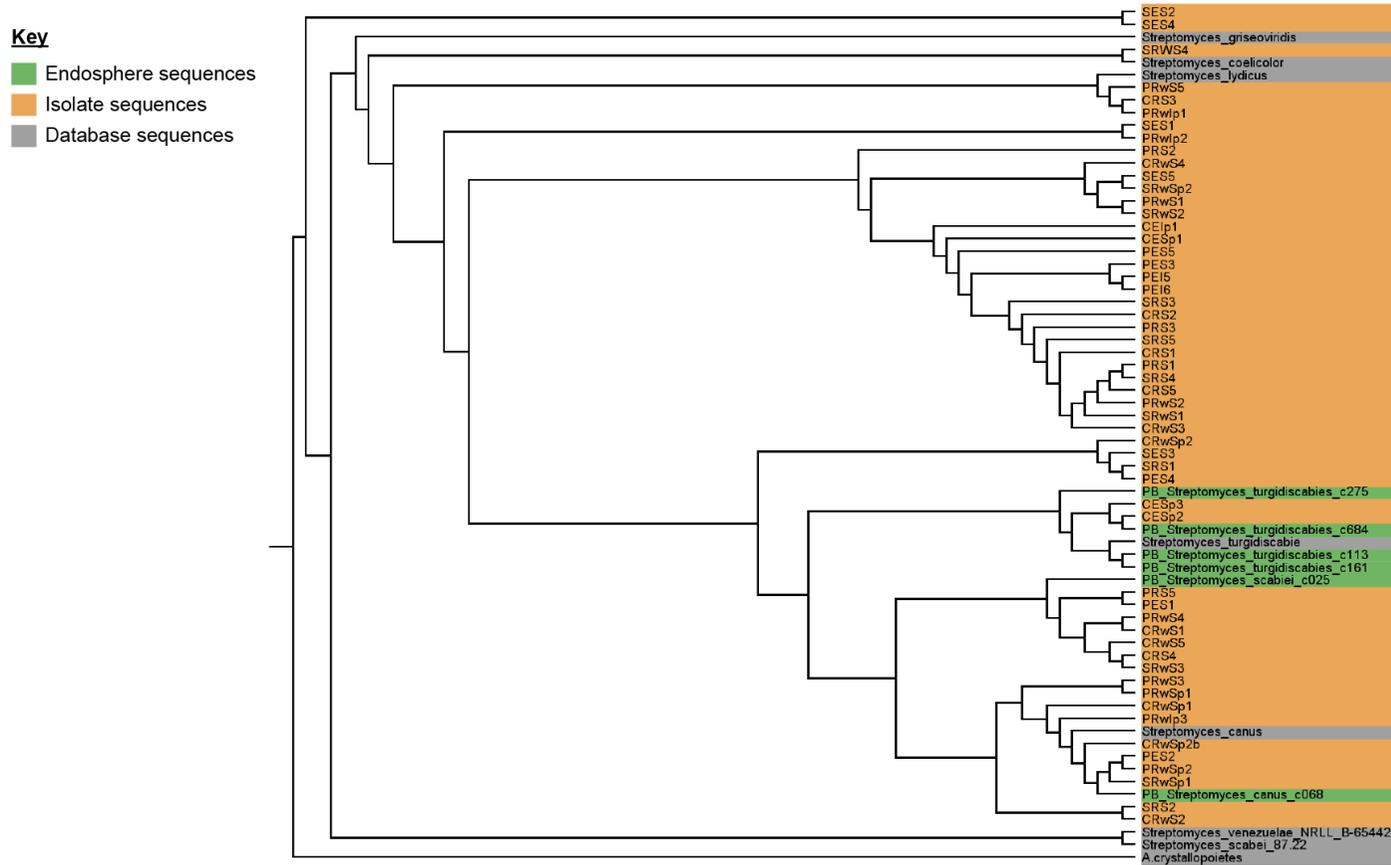


Figure 4.12. Phylogenetic tree showing the relatedness of *Streptomycetaceae* family 16S rRNA gene sequences recovered from the wheat root endosphere (green) compared to *Streptomycetaceae* root isolates of the same genera (orange), and to example 16S rRNA gene sequences acquired from the NCBI database (grey). The tree was rooted using the 16S rRNA gene sequence of a distantly related bacterium, *Arthrobacter crystallopoietes*.

# Pseudomonadaceae

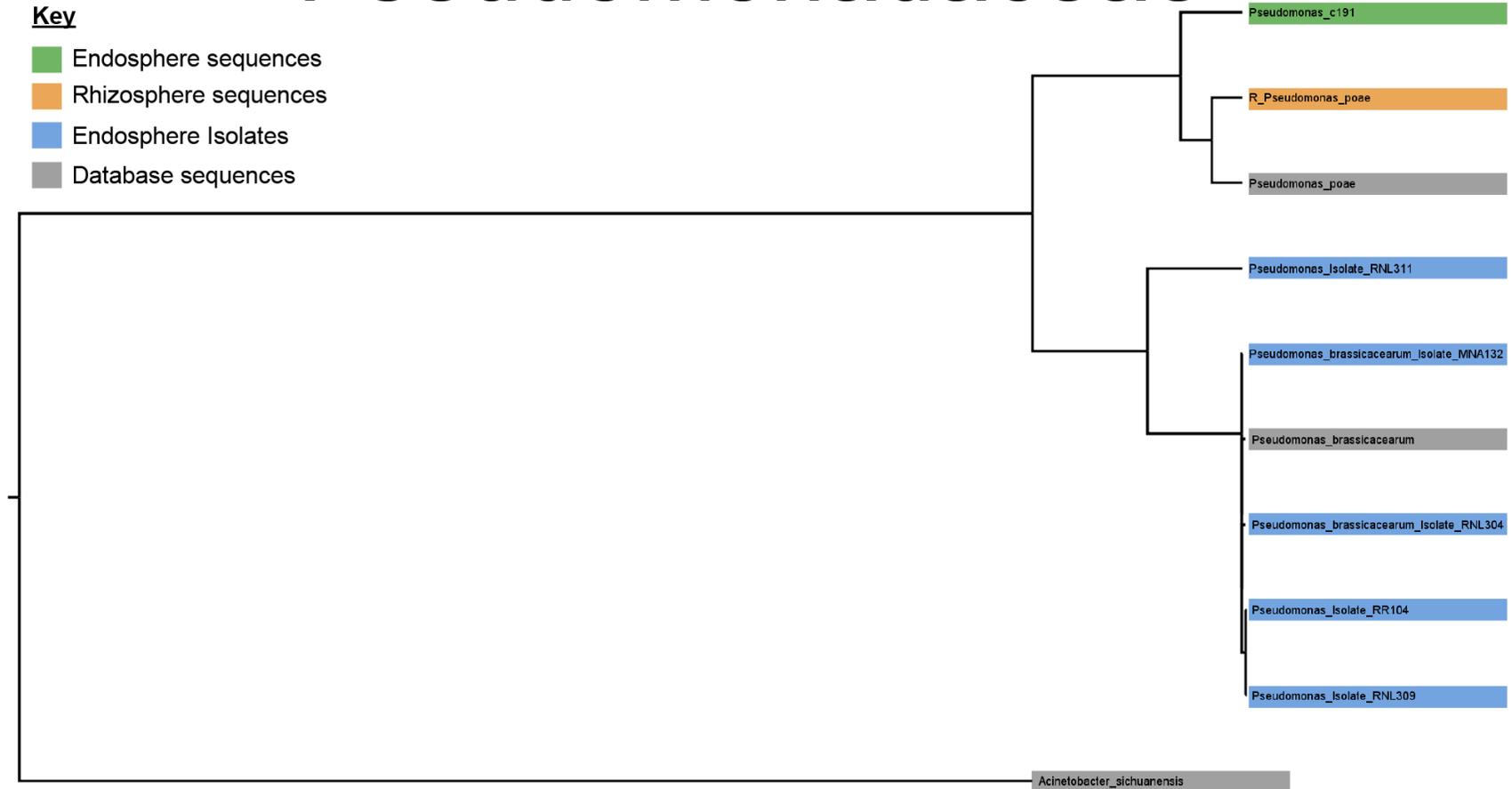


Figure 4.13. Phylogenetic tree showing the relatedness of *Pseudomonadaceae* family 16S rRNA gene sequences recovered from the wheat root endosphere (green), or rhizosphere (orange), compared to *Pseudomonadaceae* root isolates of the same genera (blue), and to example 16S rRNA gene sequences acquired from the NCBI database (grey). The tree was rooted using the 16S rRNA gene sequence of a related taxa within a different family within the Pseudomonadales order, *Acinetobacter sichuanensis*.

# Rhizobiaceae

## Key

- Endosphere sequences
- Database sequences

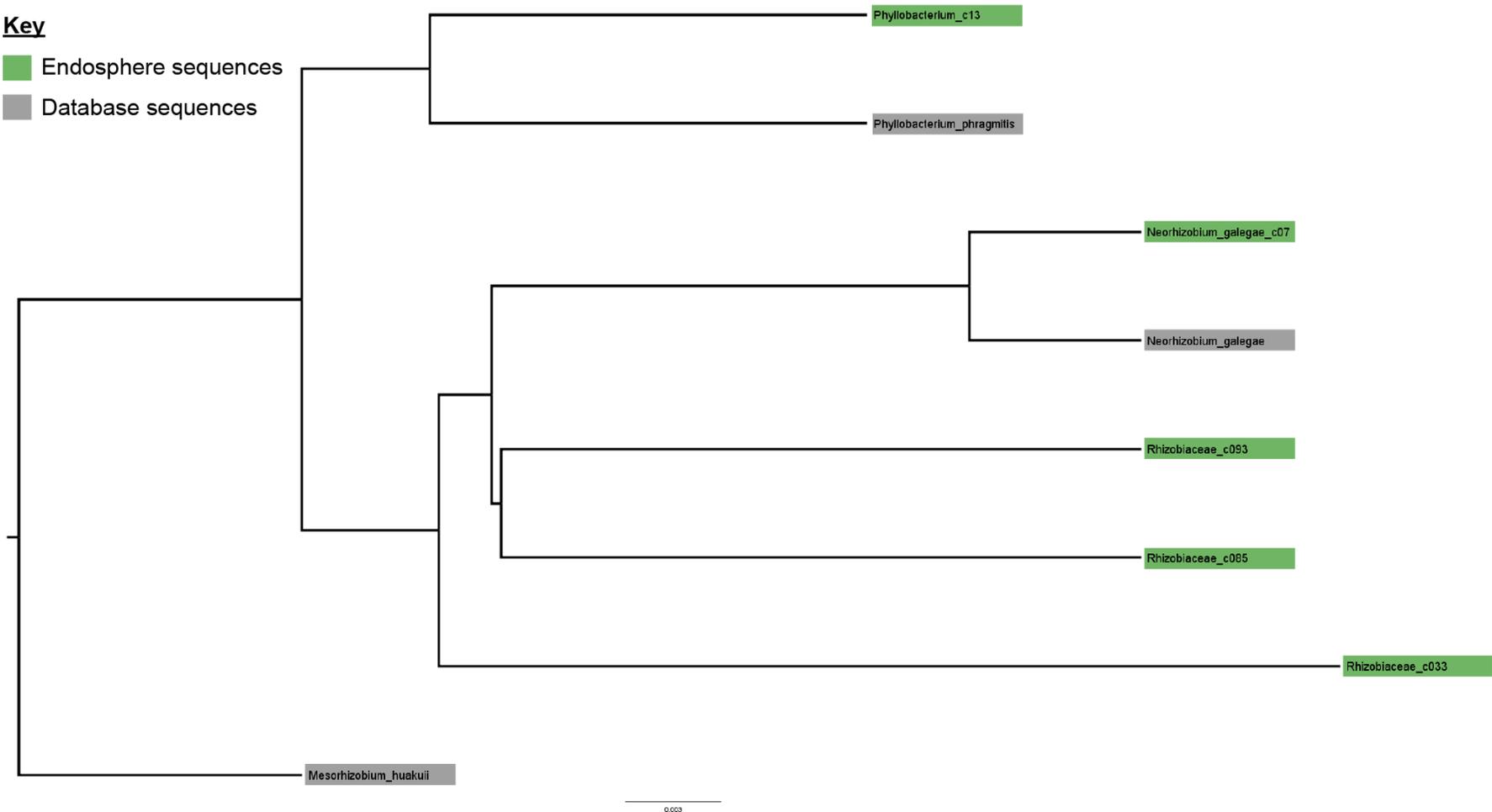


Figure 4.14. Phylogenetic tree showing the relatedness of *Rhizobiaceae* family 16S rRNA gene sequences recovered from the wheat root endosphere (green) compared to example 16S rRNA gene sequences acquired from the NCBI database (grey). The tree was rooted using the 16S rRNA gene sequence of a related taxa within a different family within the Rhizobiales order, *Mesorhizobium huakuii*.

# Chitinophagaceae

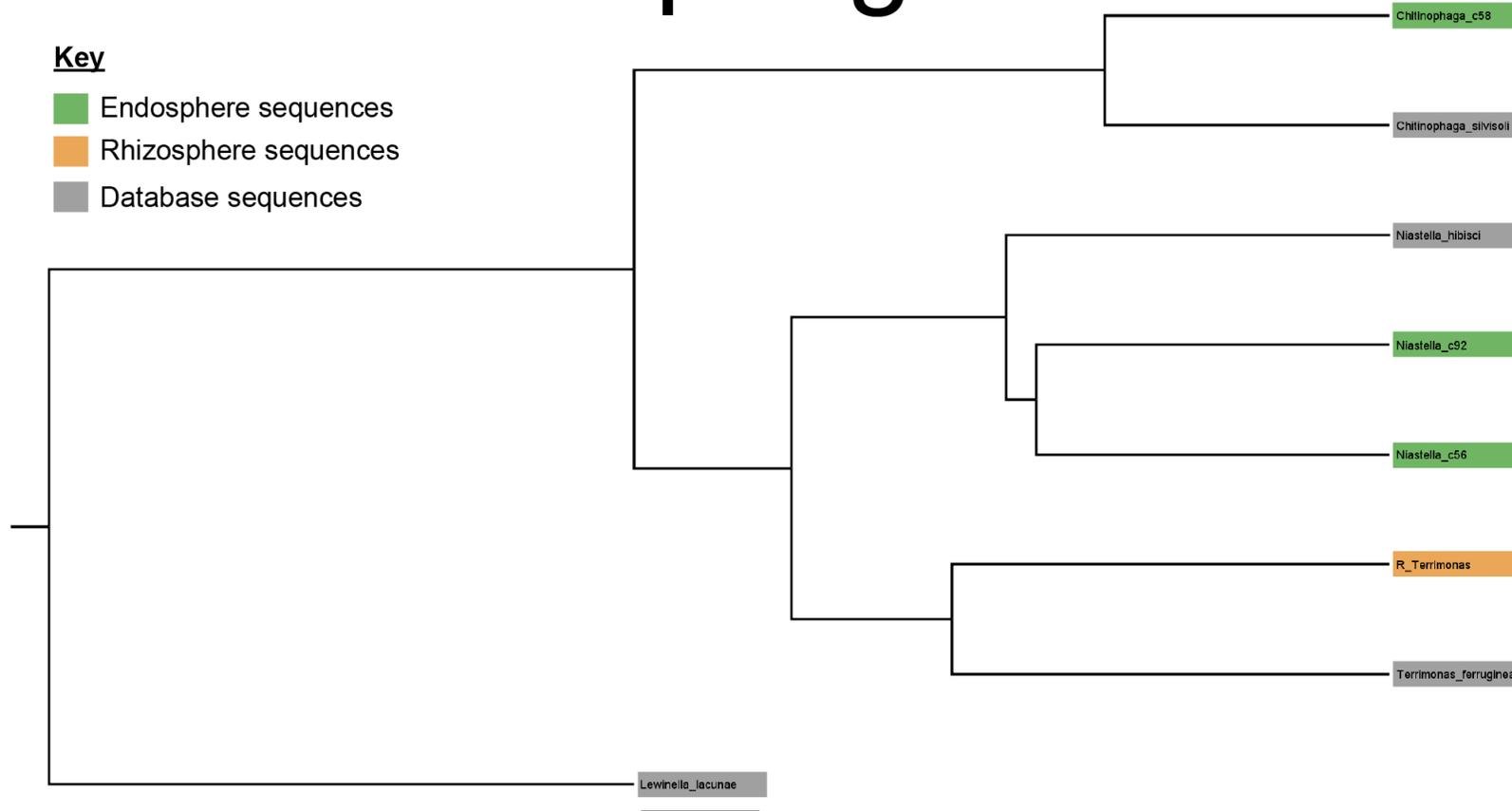


Figure 4.15. Phylogenetic tree showing the relatedness of *Chitinophagaceae* family 16S rRNA gene sequences recovered from the wheat root endosphere (green) or rhizosphere (orange), compared to example 16S rRNA gene sequences acquired from the NCBI database (grey). The tree was rooted using the 16S rRNA gene sequence of a related taxa from a different family within the order, *Lewinella lacunae*.

In addition to recovering full length 16S rRNA gene sequences for the five-core root associated taxa proposed in this chapter, sequences were also recovered for seven of the nine root exudate utilising bacteria described in chapter five. Two of these fell within the *Burkholderiaceae*, and the *Pseudomonadaceae* and *Rhizobiaceae* are also exudate utilisers. Figure 4.16 shows the phylogenetic trees associated with the three other exudate utilising families. For the *Micrococcaceae* four unique sequences were recovered. All of the endosphere sequences belonged to the *Arthrobacter* genus, in addition to one rhizosphere sequence (Figure 4.16 A) which shared >99% sequence identity with the sequence recovered from the rhizosphere, indicating these two sequences are from closely related bacteria (Supplementary Table S.14). Both of these sequences were identified as *Arthrobacter pascens* and showed >98% sequence identity with the database representative of this species, indicating that *A. pascens* may be one of the dominant *Micrococcaceae* family bacteria present within the root. One other sequence was recovered from the rhizosphere, belonging to the *Pseudarthrobacter* genus, this showed >98% sequence identity with the database representative for this taxon, strongly indicating it belongs to this genus.

Just one sequence was recovered for the *Paenibacillaceae* family (Figure 4.16 B), belonging to the *Cohnella* genus, and this group shared >96% sequence identity with the database representative for this genus (Supplementary Table S.15), indicating that this sequence belongs to the *Cohnella* genus. Lastly, for the *Cytophagaceae*, one unique sequence was recovered from the endosphere (Figure 4.16 C), belonging to the species *Cytophaga hutchinsonii*. This shared >98% sequence identity with the database representative for this species (Supplementary Table S.16), indicating that this sequence is derived from this genus. A BLAST search revealed that this sequence shared >99% sequence identity with the 16S rRNA gene of *Cytophaga aurantiaca*, so that this may be the *Cytophaga* species present within the endosphere, or the strain that is present may be closely related to both of these species.

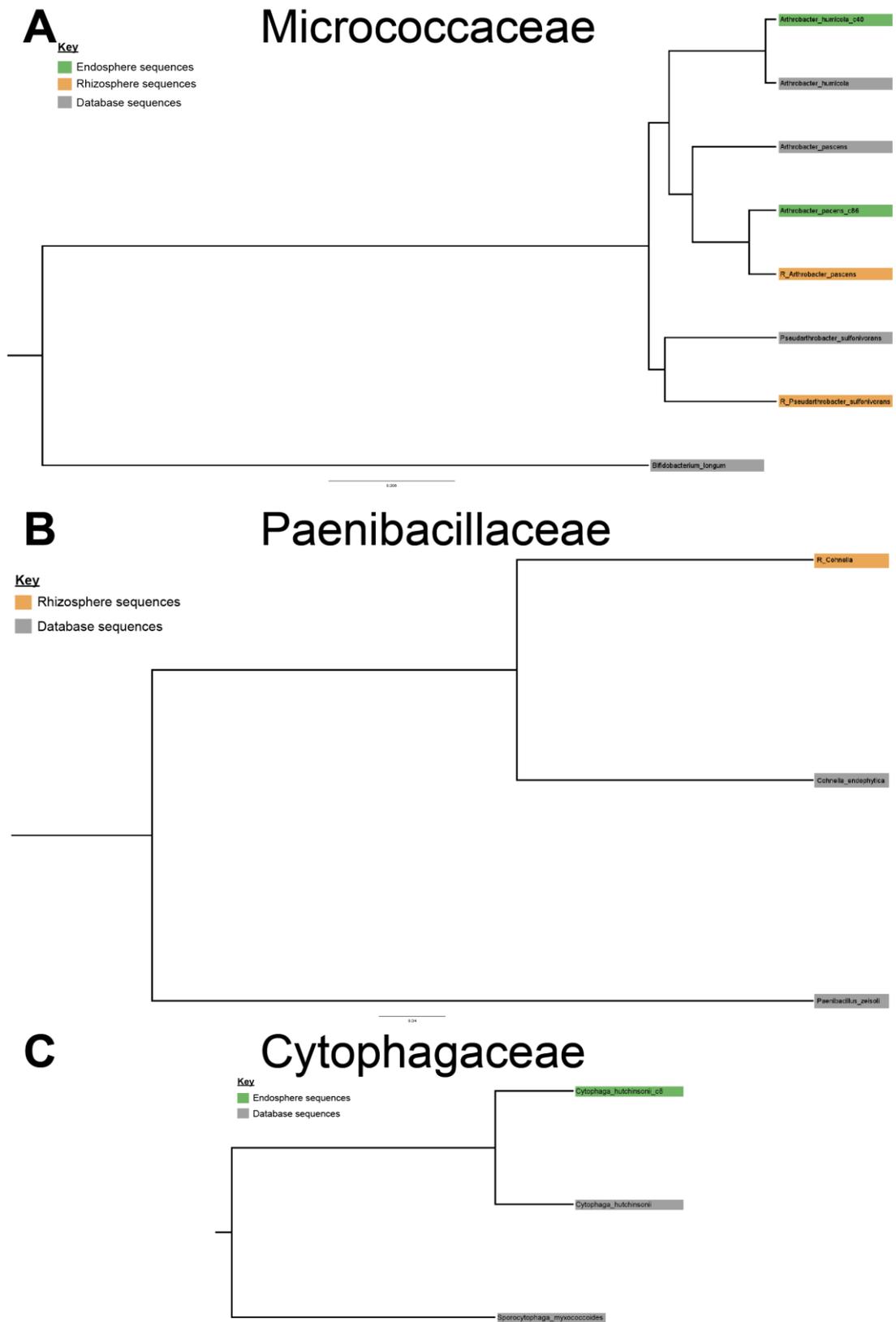


Figure 4.16. Phylogenetic tree showing the relatedness of *Micrococcaceae* (A), *Paenibacillaceae* (B), or *Cytophagaceae* (C) family 16S rRNA gene sequences recovered from the wheat root endosphere (green), or the rhizosphere (orange), compared to example 16S rRNA gene sequences acquired from the NCBI database (grey). The trees were rooted using the 16S rRNA gene sequence of a related taxa within a different family within the respective order, or a different genus within the family for each group of sequences. The rooting sequences used were from *Bifidobacterium lagum* (A), *Paenibacillus zeisoli* (B), and *Sporocytophaga myxococcoides* (C).

#### 4.6 Discussion and conclusions

This chapter has further demonstrated the abundance of *Streptomycetaceae* and *Burkholderiaceae* in the endosphere compartment of wheat roots. Five core bacterial taxa, *Burkholderiaceae*, *Chitinophageaceae*, *Pseudomonadaceae*, *Rhizobiaceae*, and *Streptomycetaceae* were enriched within the endosphere or rhizosphere of wheat regardless of soil type. When different varieties of wheat were tested, all five of these core bacterial taxa were found enriched within the endosphere or rhizosphere regardless of genotype, demonstrating that these taxa can generally associate with UK wheat varieties, and are not limited or specific to the variety Paragon. *Streptomycetaceae* and *Burkholderiaceae* were amongst the endophytes showing the greatest average relative abundance, and only the later could be detected within the seed endosphere. Longer read sequencing suggested that *Streptomyces turgidiscabies* was the dominant streptomycete within the endosphere, whereas for the *Burkholderiaceae* it was indicated that bacteria from the genera *Massilia* and *Duganella* were most prevalent.

Despite commonality, there were notable differences between the microbial communities associated with plants cultivated in agricultural soil when compared to those cultivated in Levington F2 compost, and soil was found to be a significant determinant of community composition for all three microbial groups. Most notably, the archaeal community was drastically different when comparing the two soils, whilst agricultural soil was dominated by *Nitrososphaeraceae*, the compost was dominated by *Nitrosotaleaceae*. As these are both AOA families it is likely they occupy a similar niche in each contrasting soil environment, and it is likely the different families are better adapted for each respective soil type<sup>232,250</sup>. This corroborates established literature demonstrating the impact of soil type on the wheat root microbiome<sup>56,115,117,232</sup>.

For the bacterial community, whilst present in both, groups such as *Burkholderiaceae*, *Microbacteriaceae*, *Rhizobiaceae*, *Rhodanobacteraceae*, and *Soilrubrobacterales* showed a greater average relative abundance which in the endosphere of compost cultivated plants. In agricultural soil whilst *Burkholderiaceae* were still amongst the most abundant, the groups *Chitinophageaceae*, *Paenabacillaceae*, *Streptomycetaceae*, and *Polyangiaceae* all colonised better from this soil. When the two soils were mixed, *Streptomycetaceae* showed a much greater average relative abundance of ~40%. This increased colonisation may have been the result of adding the compost to the agricultural soil; termed bio-organic

fertiliser application, co-inoculating a streptomycete biocontrol strain with organic fertiliser has been proposed as an agricultural practice as doing so can boost root colonisation and biocontrol efficacy<sup>39,340</sup>. Levington F2 compost is high in ammonia and organic matter (Table 2.3). Thus, by adding compost to agricultural soil it may have acted to boost colonisation from native *Streptomyces* species within the agricultural soil, inadvertently mimicking bio-organic fertiliser application. Intriguingly this could imply that simply adding organic fertilisers to the soil, without inoculating a non-native biocontrol strain, may be sufficient to increase colonisation by native beneficial *Streptomyces* species, where such strains are present. Previous studies demonstrating bio-organic fertiliser application showed far worse biocontrol efficacy for the uninoculated control plants however, where plants only received organic fertiliser treatment<sup>340,341</sup>. These studies however did not aim to account for the native soil community, it could be that had amenable *Streptomyces* strains been native to those soils this biostimulatory bioorganic fertiliser intervention may have been sufficient to prevent disease and promote plant health.

Nonetheless, regardless of soil inoculum seven core microbial taxa were enriched within the root associated community, five core bacterial taxa (*Burkholderiaceae*, *Chitinophageaceae*, *Pseudomonadaceae*, *Rhizobiaceae*, and *Streptomycetaceae*). Further, these five bacterial taxa were found enriched within the endosphere across all five wheat varieties tested, indicating that these taxa are core microbes for UK wheat. Few studies have investigated the core wheat root associated microbiome; amongst the studies that have been conducted a contrasting range of soils, wheat varieties, developmental timepoints, and plant cultivation practices were used<sup>97,108,109,128</sup>. Multiple potentially confounding factors makes it difficult to draw direct comparisons between individual studies. For example, Schlatter *et al.* identified *Oxolabacteraceae*, *Comamonadaceae*, *Rhizobiales*, and *Chitinophaga* as core rhizobacteria for *Triticum aestivum* cultivar Louise<sup>109</sup>. The present work corroborates this observation for multiple *T. aestivum* varieties; both *Oxolabacteraceae* and *Comamonadaceae* are now classified within the *Burkholderiaceae* (GTDB<sup>304</sup>), the Rhizobiales group contains the *Rhizobiaceae* family, and *Chitinophaga* is a genus within the *Chitinophageaceae*. Thus, these groups identified by Schlatter *et al.* are accounted for within the core wheat microbiome defined by this work. However, many of the core taxa identified by Schlatter *et al.* have not been identified by this work. It could be that these taxa are specific to the variety of wheat used by Schlatter and colleagues.

Similarly, for the endosphere community Kuźniar and colleagues identified *Flavobacterium*, *Janthinobacterium*, and *Pseudomonas* as core microbiota for both cultivars tested, and *Paenibacillus* as a core taxon for *T. aestivum* cultivar Hondia<sup>108</sup>. Of these only *Pseudomonadaceae* were identified as core endosphere microbiome members by this work, though in agricultural soil *Paenibacillaceae* were enriched within the endosphere regardless of variety (Table 4.6) and found to utilise root exudates in the rhizosphere (Chapter 5). A study by Simonin and colleagues, utilising a broad variety of wheat cultivars and soils, identified (amongst others) the bacterial taxa *Burkholderiaceae*, *Chitinophageaceae*, and *Caulobacteraceae* as core rhizosphere bacteria<sup>128</sup>, supporting the identification of *Burkholderiaceae*, *Chitinophageaceae* by the present work and by other studies<sup>109</sup>. These combined results from across multiple studies consistently imply a role for the taxa *Pseudomonadaceae*, *Rhizobiaceae*, and *Chitinophageaceae*, and for members of the *Burkholderiaceae* family within the root associated microbiota of wheat, strongly supporting the hypothesis that these are core root associated taxa for wheat. Unexplained differences however remain with some studies finding that other taxa are core microbiome components, in particular *Streptomycetaceae*. While it is likely this is largely driven by soil type, there is some evidence that for wheat, similarly to barley<sup>51</sup>, that genotype could drive differences in the root associated microbiome. Different selections of wheat cultivars were used by all of these studies, and so plant genotype may be responsible for these differences<sup>97,108,111,112</sup>. Further, in a study which used the same Church Farm field site as this work, *T. aestivum* var. Paragon was reported to be an outlier compared to other wheat varieties, with a particularly distinct rhizosphere and endosphere community<sup>97</sup>.

Contrary to this finding, the results presented in section 4.3 shows that variety was not a major factor driving community composition, and no major selection of bacterial taxa across different wheat varieties was observed. This observation supports some published results<sup>128,129</sup>. Simonin *et al.*<sup>128</sup> for example found that wheat variety had no major effect on microbial community composition. Of the published studies characterising the wheat root microbiome, taken individually the paper from Simonin and colleagues is arguably the most comprehensive as this study used a wide variety of soils with varying pH, sampled from both the African and European continent. They also used a large number of bread wheat varieties. In the present work (similarly to the findings from Simonin *et al.*) minor differences between varieties were observed, with the quantities of groups such as *Bacillaceae*, *Nocardiaceae*, or *Sphingobacteriaceae* varying significantly within the endosphere

or rhizosphere of the five different varieties tested. The major core bacterial taxa however were enriched within the endosphere and rhizosphere for all varieties.

To explain why some studies have found an effect of genotype and some have not, it could be postulated that short-read sequencing is responsible for the incongruence between the present work and these published papers. As the work presented here used short read sequencing, and thus is only able to identify microbial taxa to the family level<sup>132</sup>, changes in the abundance of individual species within those families may go undetected. Indeed this could be a caveat for all of the papers which found genotype to have no effect on the root community composition<sup>128,129</sup>. The studies which did conclude that genotype significantly influenced the community however also used short read sequencing<sup>97,108,109</sup>, so whilst this is a caveat for all of the papers being discussed and of the present work, is it unlikely to be responsible for the different conclusions drawn by these papers. As discussed in chapter one (Chapter 1 Section 1.2.2.2) within some studies which found a significant effect of genotype on root microbiome composition, this was driven by changes in the abundance of rare or low abundance taxa, which may be an unreliable measure for community dynamics<sup>105,106,316</sup>. Prior to analysis of the data comparing wheat genotypes which was presented here, prevalence-based filtering was used to remove low abundance or rare taxa likely to be sequencing artefacts<sup>316</sup>. This methodological difference could explain why the present work showed little to no effect of genotype on community composition, whereas published studies have.

Within the study conducted by Tkacz *et al.*<sup>97</sup>, where Paragon var. wheat was found to be an outlier with a distinct root microbial community, there is one other key methodological difference when compared to the present work and most other microbiome studies. Plants cultivated for this experiment were pregerminated in sterile medium before transfer to church farm agricultural soil. This technique is used in other published studies<sup>342</sup> to increase germination efficiency and to make the experimental set-up easier. The early stages of growth however, germination and initial root outgrowth, are thought to be key stages in root microbiome assembly, exemplified by the fact that root exudation profiles change significantly as the plant develops<sup>49,140</sup>. As the plant germinates and during the initial root outgrowth phase exudation levels are high and compounds are non-specific<sup>49</sup>. Microbes are recruited initially from the soil, and then the exudation profile changes across developmental stages. The community is refined via more specific exudates and antimicrobial compounds as the plant matures, to select for the core root taxa

<sup>140,343,344</sup>. For example *Arabidopsis thaliana* plants exude fewer sugars as they age<sup>49</sup>; sugars are a non-specific resource used to establish the root community in early life, then more specific compounds are secreted by mature plants to refine and maintain a specific subset of the rhizosphere community. The loss of this crucial stage through *in-vitro* pre-germination will significantly disrupt the natural microbiota recruitment process, and may explain the contrasting results between the present study, some published work<sup>128</sup>, and that of Tkacz *et al*<sup>97</sup>.

Interestingly, four of the five core enriched root associated microbial taxa were detected within the seeds, *Burkholderiaceae*, *Chitinophageaceae*, *Pseudomonadaceae* and *Rhizobiaceae* (Figure 4.10). Filamentous Gram-positive endosphere bacteria such as *Streptomyetaceae*, *Pseudonocardiaceae*, or *Nocardiaceae* were not detected within the seeds. As this experiment used short read amplicon sequencing and attempts at long read amplicon sequencing from the seed endosphere failed, it is not possible to identify the four core enriched taxa that were present within the seeds beyond the family level. Thus, we do not know if the strains of these four taxa present within the seeds are the same as those which colonised the roots. Chapter 4.2 showed that the soil had a strong influence on the root associated community, as both experiments used the same seeds this proves that the root community cannot be acquired entirely from the seed endosphere. For example, many community constituents, notably core taxa *Streptomyetaceae* and the fungal endosphere-enriched taxon *Leotiaceae* were not detected within the seeds and therefore must be recruited to the roots from the soil. It is possible that the bulk of the community is acquired from the soil, while a few core taxa such as *Burkholderiaceae*, *Chitinophageaceae*, *Pseudomonadaceae*, and *Rhizobiaceae* colonise from the seed. Longer-read amplicon sequencing or metagenomics would be required to understand if this is indeed the case.

Most fungal taxa within the seeds were classified to high taxonomic levels, such as Agaricales (an order) or Ascomycota (a phylum), making it difficult to speculate on ecological functions. The *Cladosporium* genus, which was present in all three replicates, has been noted previously as a grain associated microorganism<sup>345</sup>. This genus includes wheat pathogenic fungi such as *Cladosporium herbarum*<sup>346</sup> and the leaf mould fungus *Cladosporium fulvum*<sup>347</sup>. The *Curvularia* genus was the most dominant group within the seed endosphere for all three replicates, a genus containing wheat-pathogenic fungi that are associated with seeds<sup>348</sup>. Given the strong association of these taxa with pathogenic fungi, it is likely these genera are horizontally transmitted, seed-borne pathogens of wheat.

Little is known about how bacteria and fungi colonise the seed endosphere; though there is some evidence that microbes may migrate from the roots to the aerial plant tissues to colonise the seed <sup>339</sup> (possibly via the vascular tissue <sup>349</sup>). During developmental senescence important resources such as nitrogen are redirected from peripheral plant tissues into the developing seed <sup>223</sup>. It is possible that this redirection of plant metabolites toward the aerial plant tissues encourages root-endosphere microbes to migrate through plant vascular tissues and colonise the developing seed. This is supported by the fact that many endosphere microbiota demonstrate chemotaxis towards root exudate compounds <sup>155,157,343,350</sup>, and can traverse the plant vascular tissue <sup>210</sup>. To fully explore this hypothesis a detailed investigation is warranted combining amplicon sequencing, qPCR, and fluorescence microscopy. However, this is not the only route by which microbes could gain access to the developing seed. Another hypothesis is that microbes may colonise the seed endosphere from the anthosphere <sup>339</sup>, which refers to the microbes associated with the flowers/reproductive organs and surrounding regions. The anthosphere can be influenced by a broad range of factors including pollinators and other insects, wind-borne microbes, weather, and climatic conditions <sup>351</sup>. If the anthosphere is indeed a determinant of the seed bacterial community composition, this range of influencing factors could explain the observed stochasticity of this community. Another possibility is that microbes colonise during seed dispersal <sup>339</sup>, or in this case post-harvest processing and storage of seeds. Fumigation practices for example can significantly alter the wheat seed-associated microbiome <sup>352</sup>, though this study does not make it clear if these community shifts are driven by changes in the seed-surface microbiota or the seed endosphere microbiota. In the present work, the inconsistency of the bacterial community across replicates indicates that, irrespective of the mechanism, this process is not controlled and is subject to stochastic change. It is likely that a combination of the mechanisms discussed contributes to seed endosphere microbiome formation.

Long-read amplicon sequencing for the bacterial community was able to provide genus to species level identification for all five core bacterial families. For the *Burkholderiaceae* the genera *Massilia* and *Duganella* were found to be dominant within the endosphere, and an isolate from the *Massilia* genus was acquired from the root endosphere (Chapter 6). Both of these genera have been shown to respond positively to chitin amendments, which are performed to increase disease suppressiveness of soils <sup>353</sup>, indicating a possible role for these taxa in disease suppression. *Massilia* have been previously identified as important constituents of

the root associated microbiome as they can colonise the root in high numbers <sup>354</sup>, and have been associated with the suppression of *Rhizoctonia solani* within the wheat root microbiome <sup>355</sup>. A rhizosphere isolate from this genus was also shown to degrade cellulose, indicating occupation of plant-associated niches <sup>356</sup>. The most abundant *Burkholderiaceae* genus was *Duganella*, a genus that has shown plant growth promoting properties for wheat <sup>357</sup>, and has displayed a range of plant-beneficial activities such as phosphorous solubilisation, siderophore production, and possibly nitrogen fixation <sup>358</sup>. Further, rhizosphere inhabiting *Dugnaella* have shown antagonism toward plant pathogens <sup>359</sup>, demonstrating the potential within this genus for plant-beneficial activity.

*Massilia* are thought to be copiotrophic members of the root microbiome, particularly sensitive to nutrient availability <sup>354</sup>. As plants subjected to long read amplicon sequencing were all cultivated in agricultural soil, it is unclear if nutrient availability influences the success of *Massilia* within the root of wheat; in the more nutrient rich Levingtons F2 compost however *Burkholderiaceae* comprised 24% of the community, compared to just 15% of the community within plants cultivated in agricultural soil. If *Massilia* are the genus primarily or mostly responsible for the dominance of the *Burkholderiaceae* family then the increased relative success of this taxa in nutrient rich compost could be explained by the copiotrophic behaviour of this genera. Further to this, it has been proposed that *Massilia* can play a role phosphate cycling <sup>360</sup>. When the chemical properties for the soils used here were analysed, a SAC rating was given for key nutrients; the SAC scale is a descriptive scale for the abundance of agriculturally relevant nutrients within soils delivered by the soil analysis centre (SAC) at the James Hutton Institute (Aberdeenshire, UK). Within the agricultural soil the rating for phosphorous was 'high', at 82 mg / kg. Levington's F2 compost however was given a rating of 'extremely high' at 880 mg / kg (Chapter 2, Table 2.3). As *Burkholderiaceae* colonised the compost cultivated plants to a greater relative extent, this demonstrates a loose positive correlation between the relative abundance of *Burkholderiaceae* and both phosphorus concentrations and nutrient levels generally, adding weight to the hypothesis that the colonisation of *Burkholderiaceae* is primarily driven by copiotrophic *Massilia*, and thus that this is dependent on nutrient levels. Other factors such as pH and soil structure confound this observation however, and it remains unclear if the dominant *Burkholdereaceae* family genus within the Levington compost cultivated plants was *Massilia*, or if they primarily colonised plants grown within the agricultural soil.

For *Streptomycetaceae* the dominant species within endosphere was *S. turgidiscabies*, and these sequences showed 100% identity with two wheat endosphere isolates. These isolates were acquired by previous work within the Hutchings laboratory, and the isolations were performed from wheat plants cultivated at the same site as was used by this work. The identification of *S. turgidiscabies* as the most prominent species within the most prominent family within the endosphere of healthy wheat plants is surprising as this species is widely known as a plant pathogen, and is one of the causative agents of the disease common scab in potato (*Solanum tuberosum*)<sup>361</sup>. In addition, one other *Streptomycetaceae* sequence was annotated as *S. scabiei*, which is also known as a causative agent of common scab<sup>311</sup>, though this sequence did not cluster with the database acquired example sequence for this species (Figure 4.12). A number of the sequences were annotated as *S. canus*, a species most well-known for one isolate which was isolated from a termite associated niche, and was able to produce antifungal secondary metabolites<sup>362</sup>. This species however has also been shown to degrade phenolic compounds within soil, which are common root exudates, particularly for wheat<sup>140</sup>. *S. canus* was also shown to activate the expression of antioxidant enzymes in cucumber seedlings (*Cucumis sativus*), which reduces reactive oxygen stress within the leaves<sup>363</sup>. *S. canus* has been isolated from the mycorrhizal fungus *Glomus mosseae*, and shown an ability to both promote the growth of pomegranate (*Punica granatum*), and provide some protection against disease<sup>364</sup>. Thus, for *S. canus* there is precedent in the literature for this species to be a plant beneficial.

The most abundant *Streptomycetaceae* sequences were all assigned to plant-pathogenic *Streptomyces*, *S. turgidiscabies* and *S. scabiei*. This observation is surprising given that the plants that were sampled were healthy and showed no sign of disease. One hypothesis that could explain this observation is that these beneficial or neutral streptomycetes have evolved from or have evolved into the plant pathogen *S. turgidiscabies*, and thus share a closely related 16S rRNA gene sequence. In order to adapt from a pathogenic to a mutualist lifestyle, these species could have lost the pathogenicity islands, which contain genes such as *txtAB*<sup>314</sup> used to cause disease, but maintained the ability to colonise the root tissue. Conversely plant beneficial *Streptomyces* spp. could have gained this pathogenicity island and thus adopted a pathogenic lifestyle. The data presented in chapter three showed that *Streptomycetaceae* are unable to persist within the roots of wheat after the plants senesced and the tissue began to die. This implies that if the *S.*

*turgidiscabies* residing within the endosphere were to kill the plant through pathogenesis, or if they were to cause tissue necrosis, then they would quickly be outcompeted within that niche by necrotrophic fungi colonising the dead plant tissue. This supports the hypothesis that the *S. turgidiscabies* within the wheat endosphere would benefit from exhibiting a less severe disease phenotype, to maintain tissue integrity and protect themselves from niche invasion by necrotrophic fungi. In such a scenario, where living plant tissue is the required niche for *S. turgidiscabies*, it would become beneficial for this strain to be less virulent, and eventually to become a beneficial member of the microbial community. Further, it may aid the proliferation of the strain to adapt to promote plant health and growth, thus maximising the size of its own endospheric niche. One example supporting this hypothesis is of a *S. scabiei* isolate that was found to inhibit the wheat take all fungus, and the culture filtrate for which suppressed infection *in planta*. The authors also show how this subspecies of the *S. scabiei* was avirulent, demonstrating that *Streptomyces* strains within typically pathogenic species have the capacity to lose the ability to infect plant tissue, and then to demonstrate plant-beneficial traits <sup>365</sup>. This is one possible explanation then for how *S. turgidiscabies* can be prevalent within the roots of healthy plants, though the present work can present no direct evidence for this.

Compared to many other crop diseases, which can cause severe yield losses or even total loss of the crop, common scab is not often a severe disease, with the bacteria only infecting the outermost cell layers <sup>314</sup>. Thus, the causative agents of common scab are able to be successful without causing major damage to the host plant. One other explanation, if the dominant streptomycetes within the endosphere are indeed *S. turgidiscabies*, is that the that the severity of disease caused by these strains is low enough within wheat roots to not cause significant stress to the host plant, or to cause the major tissue necrosis seen within potato common scab. These strains may have adapted in this way to take advantage of host resources without causing significant damage and tissue necrosis, as a mechanism to protect the endosphere niche from colonisation by necrotrophs.

The most important consideration for the interpretation of this data is the efficacy of 16S rRNA gene sequences for the identification of species, particularly within the *Streptomyces* genus. For Actinobacteria in particular (the phylum which contains *Streptomyces*) 16S rRNA gene sequences have been reported to show limited capacity for revealing phylogenetic relationships <sup>366</sup>. *Streptomyces* species can be particularly hard to distinguish using 16S rRNA gene sequences alone, for example

it has been shown that different *Streptomyces* species can share identical 16S rRNA gene sequences <sup>367</sup>, and these bacteria are known to possess multiple copies of the gene. This complicates the resolution of *Streptomyces* species via the 16S rRNA gene; whilst we can be confident all of the *Streptomycetaceae* endosphere sequences belong to *Streptomyces*, it is possible that they are derived from different species even whilst sharing 100% sequence identity. The most likely explanation then is that the *Streptomyces* from which these species originated may be closely related to *S. turgidiscabies*, but there is no evidence to suggest they might share the same biological capabilities or ecological strategy. To address this limitation an alternative marker gene could be used, such as any of the Actinomycete specific genes proposed by Gao and colleagues <sup>366</sup>, or universally conserved genes such as *rpoB* <sup>368</sup>. Alternately a multi-locus approach could be taken, as has been used for *Streptomyces* <sup>369</sup>, though for whole community scale studies this may not be practical.

For the *Pseudomonadaceae*, despite all of root isolates from this family being of the *Pseudomonas brassicacaerum* species (Chapter 6), the sequences recovered from the RAM were *P. poae*. This pseudomonad species is commonly isolated from plant root endosphere compartments <sup>370,371</sup>, and can provide a range of services to the host including disease suppression <sup>370,372</sup>, plant growth promotion <sup>371,372</sup>, phosphate solubilisation <sup>373</sup>, and mycotoxin degradation <sup>372</sup>. This species has also demonstrated beneficial capabilities within the wheat RAM <sup>373</sup>, and has been identified within the phyllosphere of wheat <sup>374</sup>. *P. poae* shows precedent as a plant beneficial microbe and is likely to fulfil this role within wheat.

Within the *Rhizobiaceae* family a number of genera were detected within the rhizosphere and endosphere, *Rhizobium*, *Agrobacterium*, *Phyllobacterium*, and *Neorhizobium*. Specifically *Neorhizobium galegae* was identified, a species commonly associated with symbiosis and nitrogen fixation <sup>375</sup>. *Rhizobium* are also well known for nitrogen fixation, most notably within the root nodules of legumes such as pea (*Pisum sativum*) <sup>376</sup>. This genus has also demonstrated the capacity to promote the growth of some wheat varieties, though for some varieties inoculation with *Rhizobium* had little effect <sup>377</sup>. *Rhizobium* can also help support the growth and maintain yields for wheat experiencing drought stress <sup>378</sup>. *Phyllobacterium* is also a root-associated genus with plant growth promoting properties <sup>379</sup> such as phosphate solubilisation <sup>380</sup>. For wheat, this genus has been shown to solubilise phosphate within the rhizosphere at particular growth stages <sup>381</sup>. *Agrobacterium* on the other hand, which was the most abundant individual taxon, is most well-known for plant

pathogens such as *Agrobacterium tumefaciens*, and the causative agent of crown gall disease and a bacterium used widely for the transformation of plant cells <sup>382</sup>. Whilst *Rhizobium*, *Phyllobacterium*, and *Neorhizobium* all share similar traits, and are likely to be beneficial members of the community, the presence of *Agrobacterium* indicates that the presence of *Rhizobiaceae* within the roots can also be the result of infection. The diversity contained within this group implies that changes in the abundance of individual genera or species may be masked by family level taxonomic assignments within short-read amplicon studies.

*Niastella*, *Chitinophaga*, and *Terrimonas* were three *Chitinophagaceae* genera identified within the endosphere and rhizosphere. *Niastella* species have been identified and isolated from within the rhizosphere <sup>383,384</sup>, and have demonstrated the ability to both utilise host-derived carbon and to solubilise phosphorous <sup>385</sup>. *Chitinophaga* have also previously been identified within root associated microbiomes <sup>386,387</sup>, including for wheat <sup>123,388</sup>, have been associated with disease suppression <sup>388</sup>, and are known as cellulose degraders <sup>387</sup>. *Terrimonas* were identified within the rhizosphere, a genus which has been isolated from the rhizosphere in the past <sup>389,390</sup>. The most abundant of these genera was *Niastella*, though it is unclear what role these organisms may play within the wheat root, members of this genus are capable of solubilising phosphorous and utilising root exudates. The identification of these, and all of the other genera and species from the core root associated families lends direction to future studies aiming to isolate root associated microbiota and characterise the interactions between these organisms and the host.

This chapter has shown that while soil inoculum was a major determinant of the root associated microbial community, a core group of five core microbial taxa were enriched within the endosphere of wheat regardless of the soil condition or plant variety. These taxa were *Burkholderiaceae*, *Chitinophagaceae*, *Pseudomonadaceae*, *Rhizobiaceae*, and *Streptomycetaceae* and the core fungal taxa *Leotiaceae* and *Moritrellaceae*. Some of these bacterial taxa were also detected within the seed interior. Whilst no archaea were detected within the seed interior, a variety of fungal and bacterial taxa were identified. The significant influence of soil inoculum on root community composition, and the fact that a number of core root taxa (most notably *Streptomycetaceae*) were absent from the seed interior, clearly demonstrates that the root community is primarily recruited from the surrounding soil. Regardless, the presence of core root associated families within the seed, and the demonstration via qPCR that in principle bacteria can

colonise the root from the seed, raises questions surrounding how these bacteria colonise the seed, if seed-endosphere strains are the same as we observe within the root endosphere, and if these taxa are horizontally transferred between generations via the seeds. Contrary to some published studies, a surprisingly consistent bacterial community was found associated with the roots of five different wheat varieties. Particularly striking was the similarity between the closely related varieties Paragon, Cadenza, Axona and Tonic, and the more distantly related variety Soissons, indicating that genotype has a relatively weak impact on the root associated community. When provided the same starting soil community, different UK wheat varieties are able to recruit similar microbial taxa, and thus must be using similar mechanisms (for example, similar secretions of host derived compounds) to recruit the root endosphere microbiome from the soil. Overall, these results show that UK wheat varieties consistently recruit a core group of bacterial taxa to colonise the roots from the soil. Most notably, *Streptomycetaceae* and *Burkholderiaceae*, two bacterial families which are always amongst the most abundant within the endosphere compartment.

## **Chapter 5. Identification of root exudate utilising microbes within the wheat root microbiome using DNA stable isotope probing (SIP)**

Sections of this chapter have been published previously <sup>198</sup>.

### **Aim and approach**

After the identification of five core microbial taxa associated with wheat roots, we aimed to understand how these taxa are recruited to the root microbiome from the soil. 30-40% of the carbon fixed photosynthetically from the atmosphere by plants is exuded from the roots as root exudate compounds <sup>138</sup>. These compounds can be utilised as a carbon source by microbes residing within and in the vicinity of the root, and plants can tailor exudate composition to recruit and sustain specific microbial taxa within the root. Thus, we aimed to identify the microbial taxa that wheat can support via these root exudate compounds using <sup>13</sup>CO<sub>2</sub> DNA-SIP. Briefly, triplicate wheat plants were cultivated either in <sup>13</sup>CO<sub>2</sub>, or <sup>12</sup>CO<sub>2</sub> for two weeks. During this period, for plants cultivated with the “heavy” <sup>13</sup>CO<sub>2</sub>, <sup>13</sup>C becomes photosynthetically fixed and incorporated into the plant’s metabolism. Thus, carbon-based metabolites such as root exudates become labelled with <sup>13</sup>C before export from the roots. Microbial utilisation of these compounds then results in the <sup>13</sup>C label being incorporated into the DNA backbone of actively growing microorganisms within the rhizosphere or endosphere. Heavy and light DNA can be separated via density gradient ultracentrifugation and fractionation, and after this process the fractions were analysed using amplicon sequencing to identify metabolically active microbes. A two-week labelling period was chosen to minimise the probability of labelling via cross feeding by secondary metabolisers <sup>183,189</sup>. To control for CO<sub>2</sub> fixing autotrophs residing within the soil, an unplanted soil control was also incubated with <sup>13</sup>CO<sub>2</sub>, such that any microbes within the soil able to fix CO<sub>2</sub> can also be detected using the <sup>13</sup>C label and discounted from analysis of exudate utilising taxa. Using this approach, the aim of this chapter was to identify root exudate or host-derived carbon utilising fungal, archaeal, and bacterial taxa within the rhizosphere and endosphere of UK elite spring bread wheat variety, *T. aestivum* Paragon.

## Results

### 5.1 Archaeal 16S rRNA gene SIP

A number of studies have used DNA-SIP to identify root exudate utilising fungi and bacteria within both wheat<sup>65,66,188,197,391</sup> and a range of other plants<sup>161,391,392</sup>. The ability of archaea to utilise root exudates has only previously been assessed for rice<sup>191</sup>, and little is known about how archaea might interact with the roots of terrestrial plants. Thus, DGGE and qPCR were used to assess whether the archaeal community shifted towards the heavier DNA fractions, such a shift would be indicative of a subset of the archaeal community utilising root exudates<sup>189</sup>. PCR amplification of the archaeal 16S rRNA gene from endosphere SIP fractions was inconsistent (Figure 5.1 A), so DGGE could not be performed from the endosphere compartment. Inconsistent amplification could have resulted from low quantities of archaeal DNA within the endosphere fractions. Amplification from the rhizosphere however was consistent, though DGGE did not indicate a shift for the archaeal community (Figure 5.2). This was confirmed by qPCR, which showed that there was no shift towards the heavy fractions for the archaeal community within the rhizosphere, and showed this was also the case for the endosphere (Figure 5.3). Overall these results suggest that archaea were unable to utilise host-derived carbon during the two-week labelling period, and so no sequencing of the archaeal 16S rRNA gene was performed on the heavy and light DNA SIP fractions.

### 5.2 Fungal 18S ITS2 region SIP

Previous DNA-SIP studies have shown that fungi associated with the roots of wheat are able to utilise root exudates<sup>393</sup>. Most published fungal DGGE primers are not specific to the fungal ITS2 region however and will amplify *T. aestivum* ITS2 sequences. Thus, these primers are not useful for a host-associated system. PCR amplification of the fungal ITS2 region was instead performed using the fungi specific primers ITS7F/ITS4R\_2; amplification was not consistent, and no band was yielded for the majority of fractions (Figure 5.1 B C), possibly indicating low quantities of fungal DNA within the fractions. As a result, DGGE was not performed on the fungal community. qPCR also indicated that fungal DNA was not consistently distributed across all 12 fractions and did not indicate SIP-labelling of the fungal community for either the rhizosphere or endosphere (Figure 5.4). Overall these results suggest that root-associated fungi did not utilise host-derived carbon during the two-week labelling period and so no sequencing of the fungal ITS2 region was performed.

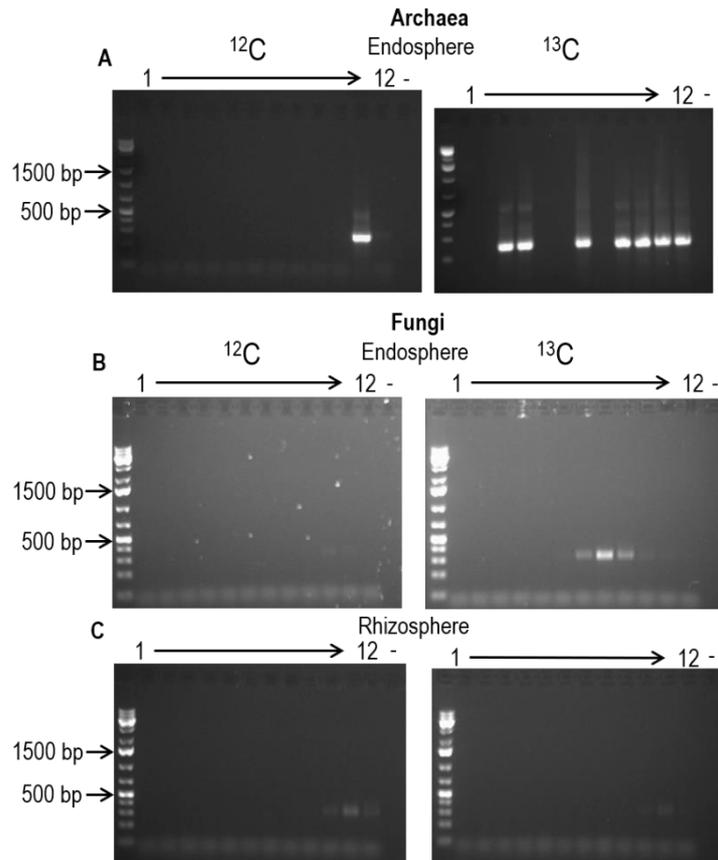


Figure 5.1 Agarose gel electrophoresis performed from SIP fractions to test the amplification of archaeal 16S rRNA gene sequences or fungal ITS2 sequences prior to DGGE. Fractions shown are from  $^{12}\text{C}$  (left) or  $^{13}\text{C}$  (right) labelled plants. **A** shows the second round of the nested PCR approach for the amplification of the archaeal 16S rRNA gene from endosphere SIP fractions (primers A109F/A1000R, expected band size 1000 bp). **B** shows the attempted amplification of the fungal ITS2 region from endosphere SIP fractions 1-12. **C** shows the attempted amplification of the fungal ITS2 region from rhizosphere SIP fractions 1-12 (primers for both fungal PCRs were fITS7F/ITS4R\_2 expected size 500bp). Lanes labelled - were negative controls.

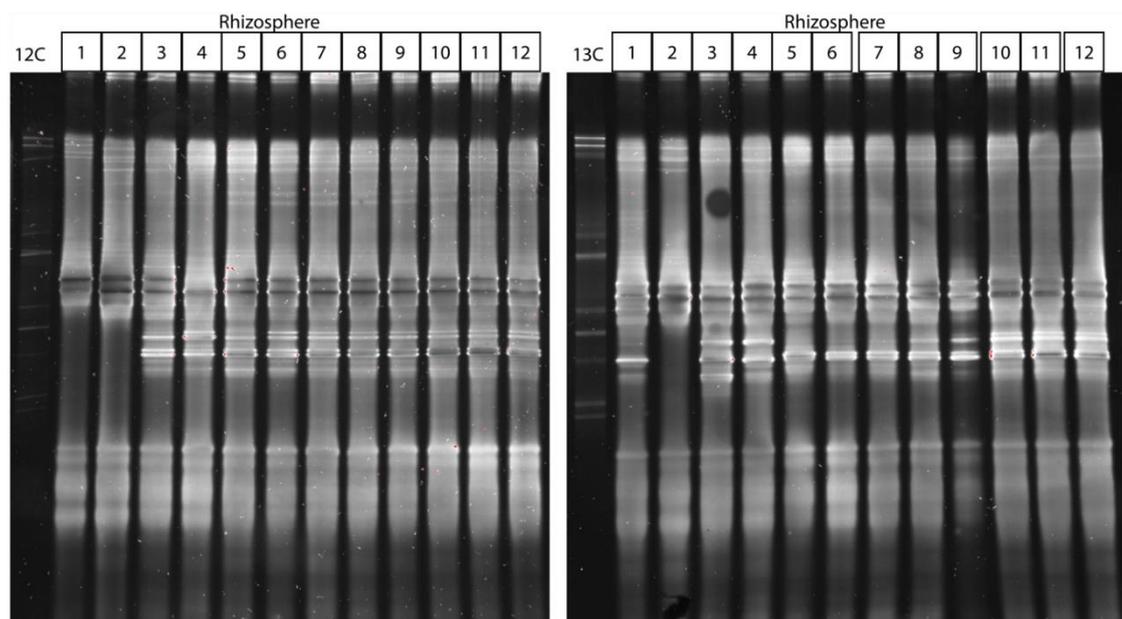


Figure 5.2 Denaturing gel gradient electrophoresis (DGGE) performed to assess the shift in the archaeal community composition across rhizosphere SIP fractions. Lanes show 16S rRNA gene diversity across the 12 SIP fractions from rhizosphere compartment for with one replicate each for  $^{12}\text{C}$  control (left) and  $^{13}\text{C}$  labelled plants.

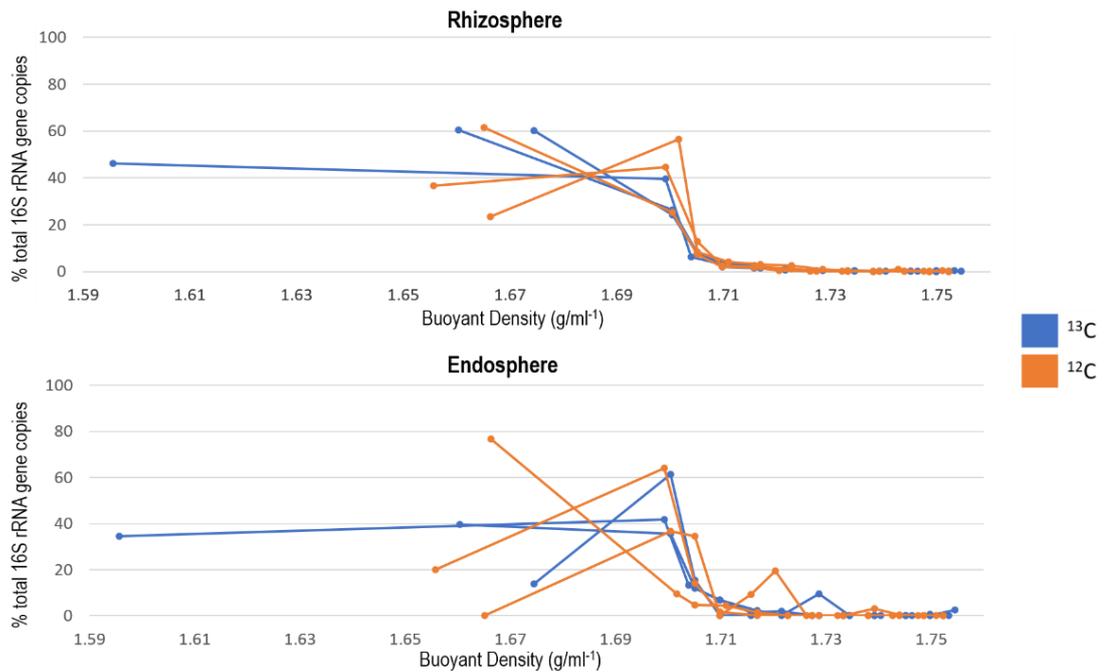


Figure 5.3 Quantitative PCR performed to assess if there had been a shift in the archaeal community toward the heavy fractions, which would be indicative of  $^{13}\text{C}$  incorporation. qPCR was performed against the archaeal 16S rRNA gene to test for  $^{13}\text{C}$  labelling of the archaeal community across fractions. Graphs shows the percent of total 16S rRNA genes found within each of the 12 fractions for each plant (plotted as buoyant densities for that fraction in  $\text{g} / \text{ml}^{-1}$ ) for  $^{12}\text{C}$  control (orange) and  $^{13}\text{C}$  labelled (blue) wheat plants from rhizosphere (top) and endosphere compartments (bottom) (N=3, each replicate is plotted individually).

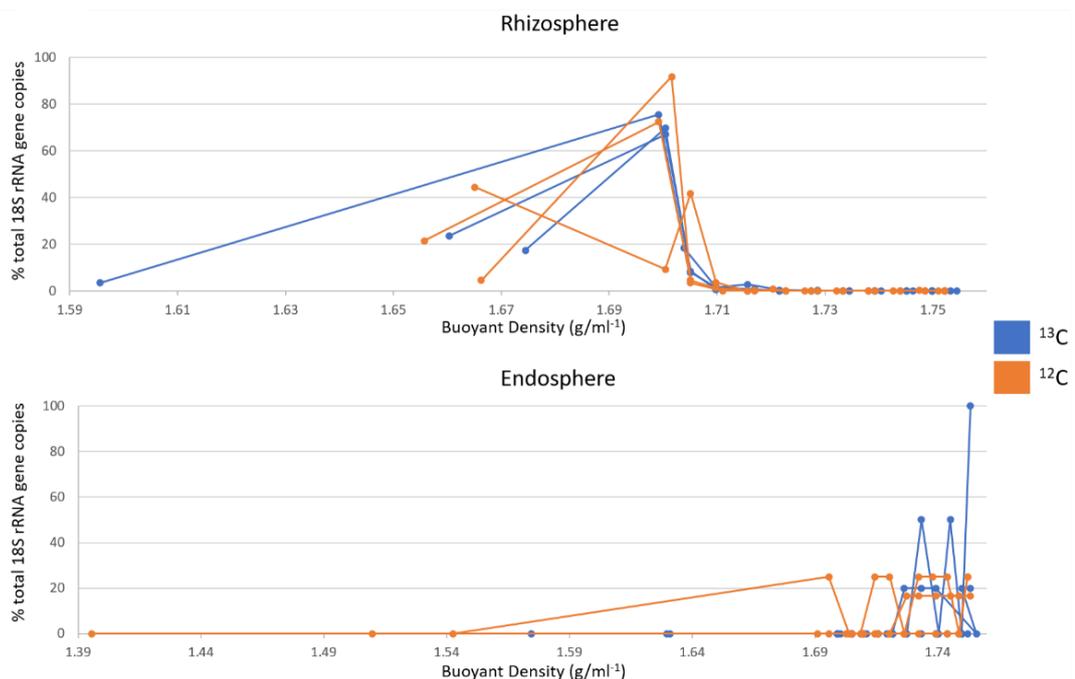


Figure 5.4 Quantitative PCR performed to assess if there had been a shift in the fungal community toward the heavy fractions, which would be indicative of  $^{13}\text{C}$  incorporation. qPCR was performed against the fungal 18S rRNA gene to test for  $^{13}\text{C}$  labelling of the fungal community across fractions. Graphs shows the percent of total 18S rRNA genes found within each of the 12 fractions for each plant (plotted as buoyant densities for that fraction in  $\text{g} / \text{ml}^{-1}$ ) for  $^{12}\text{C}$  control (orange) and  $^{13}\text{C}$  labelled (blue) wheat plants from rhizosphere (top) and endosphere compartments (bottom) (N=3, each replicate is plotted individually).

### 5.3 Bacterial 16S rRNA gene SIP

#### 5.3.1 CO<sub>2</sub> fixation by soil autotrophs

Within the unplanted bulk soil controls differential abundance analysis indicated that six bacterial taxa were significantly enriched in the heavy DNA fraction compared to the light, and so were hypothesised to fix <sup>13</sup>CO<sub>2</sub> autotrophically (Table 5.1). Only one of these taxa was also <sup>13</sup>C-labelled within the rhizosphere of <sup>13</sup>C labelled plants, *Intrasporangiaceae* (Supplementary Table S.8). This taxon was therefore excluded from the list of root exudate utilising bacterial taxa. While microbes belonging to this family are capable of photosynthesis, they also have genomes with high GC content, and as such they may be overrepresented in heavy fractions.

**Table 5.1. Differential abundance analysis for CO<sub>2</sub>-fixing autotrophs**

Differential abundance analysis performed using DESeq2 to identify significantly differentially abundant taxa heavy fractions from the <sup>13</sup>C treated unplanted soil control when compared to light fractions. This was performed to identify soil taxa that are capable of fixing <sup>13</sup>CO<sub>2</sub> from the headspace of plant cultivation chambers, such that autotrophic, phototrophic, or heterotrophic <sup>13</sup>CO<sub>2</sub> fixation can be controlled for when identifying root exudate utilisers.

Taxa	baseMean	log2FoldChange	lfsSE *	padj *
<i>Gaiellaceae</i>	523.2173795	2.423212929	0.588347809	0.000181456
<i>Gemmatimonadaceae</i>	463.0262448	2.071076321	0.535193797	0.000453911
<i>Acidimicrobiaceae</i>	366.6391772	1.760321921	0.508212808	0.001886839
<i>Micromonosporaceae</i>	152.0272389	1.923264409	0.628460704	0.005528292
<i>Solirubrobacteraceae</i>	114.7633207	1.91416848	0.571306195	0.002444309
<i>Intrasporangiaceae</i>	111.1246338	2.435159256	0.928676387	0.016801645

\*lfsSE = log2 Fold Change Standard Error, padj = P adjusted

#### 5.3.2 Endosphere compartment

DGGE indicated that the bacterial community composition shifted towards the heavy fractions for <sup>13</sup>C labelled plants when compared to the <sup>12</sup>C control (Figure 5.5). It must be noted however that the primers used for DGGE (PRK341F-GC / MPRK806R) are not specific to the bacterial community, and for amplicon sequencing experiments with these primers, taxonomy-based filtering is used to remove contaminating host derived sequences (Chapter 2, Section 2.7). It is possible then that the shift observed from this DGGE experiment is the result of host derived DNA becoming labelled, and not due to labelling of the microbial community within the endosphere. After host derived sequences had been removed from the

sequencing data (Supplementary Table S.1) PCoA indicated that bacterial communities within endosphere samples were highly variable (Figure 5.6 B), and permanova confirmed that there was no significant difference between  $^{13}\text{C}$ -labelled heavy and light fractions (permanova:  $R^2 = 0.29$ ,  $p > 0.1$ ), this is likely to be due to the high variability between samples (Figure 5.6 B), further reflected by the high variability between the unfractionated community of the three endosphere replicates (Supplementary Figure S.5). This means that endosphere dataset was too variable to draw any conclusions from the current study about the utilisation of host derived carbon within the endosphere.

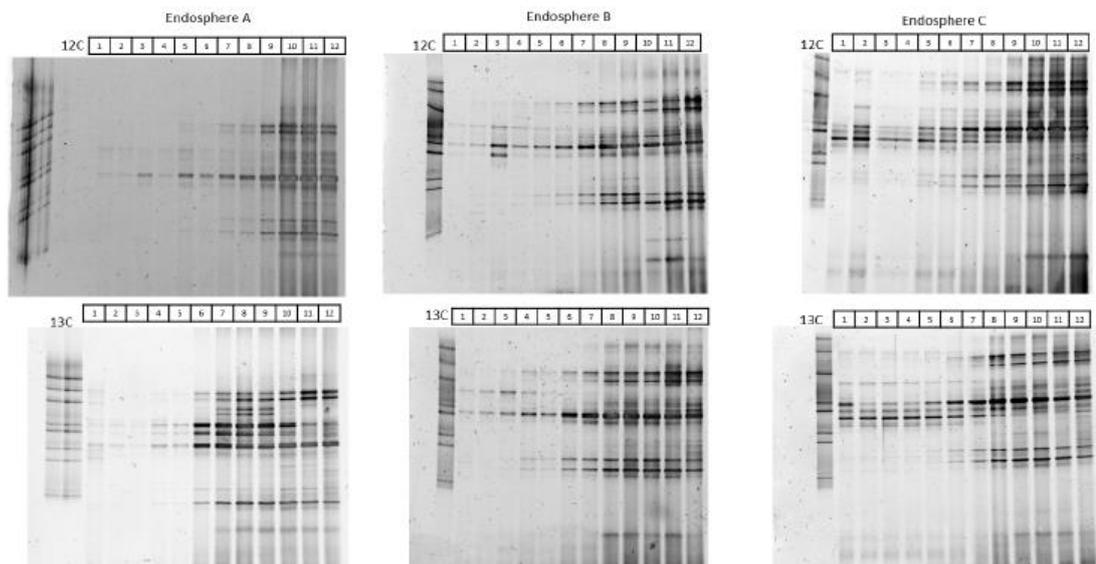


Figure 5.5 Denaturing gel gradient electrophoresis (DGGE) performed to assess the shift in the bacterial community composition across rhizosphere SIP fractions. Lanes show bacterial 16S rRNA gene diversity across the 12 fractions generated for stable isotope probing for the endosphere associated with three  $^{12}\text{C}$  control (top) and  $^{13}\text{C}$  labelled (bottom) plants. These gels show a shift in the bacterial community towards the heavy fraction of  $^{13}\text{C}$  labelled plants.

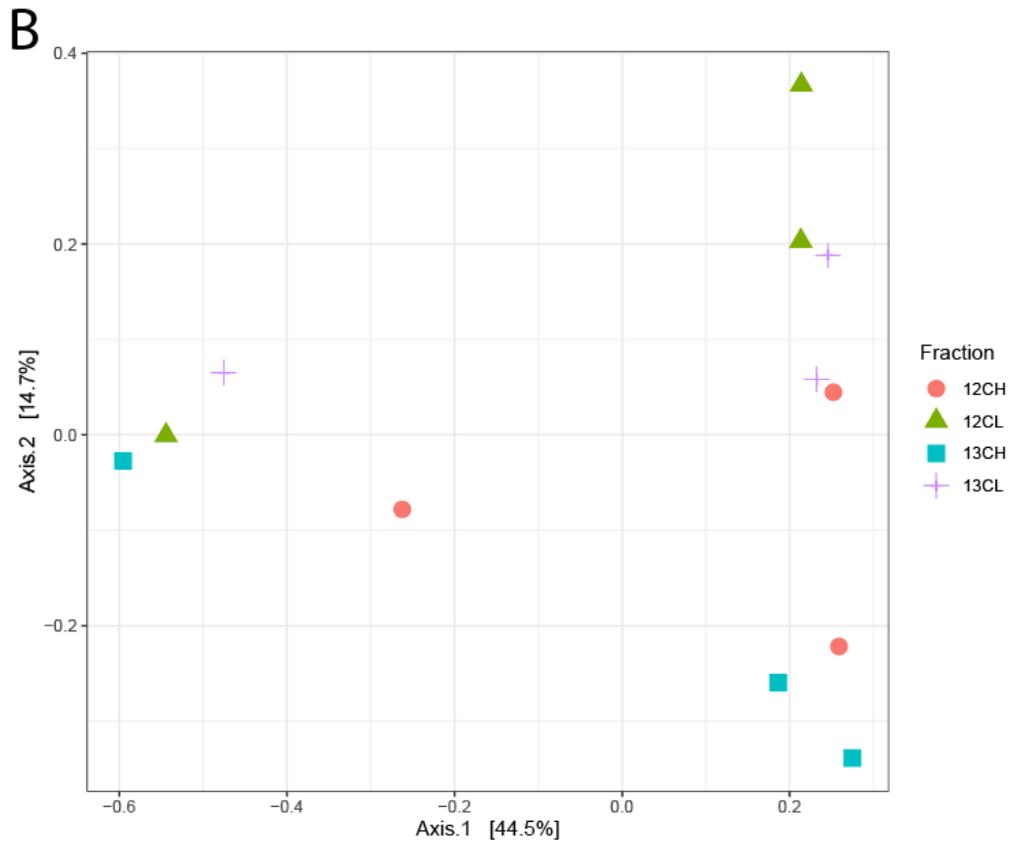
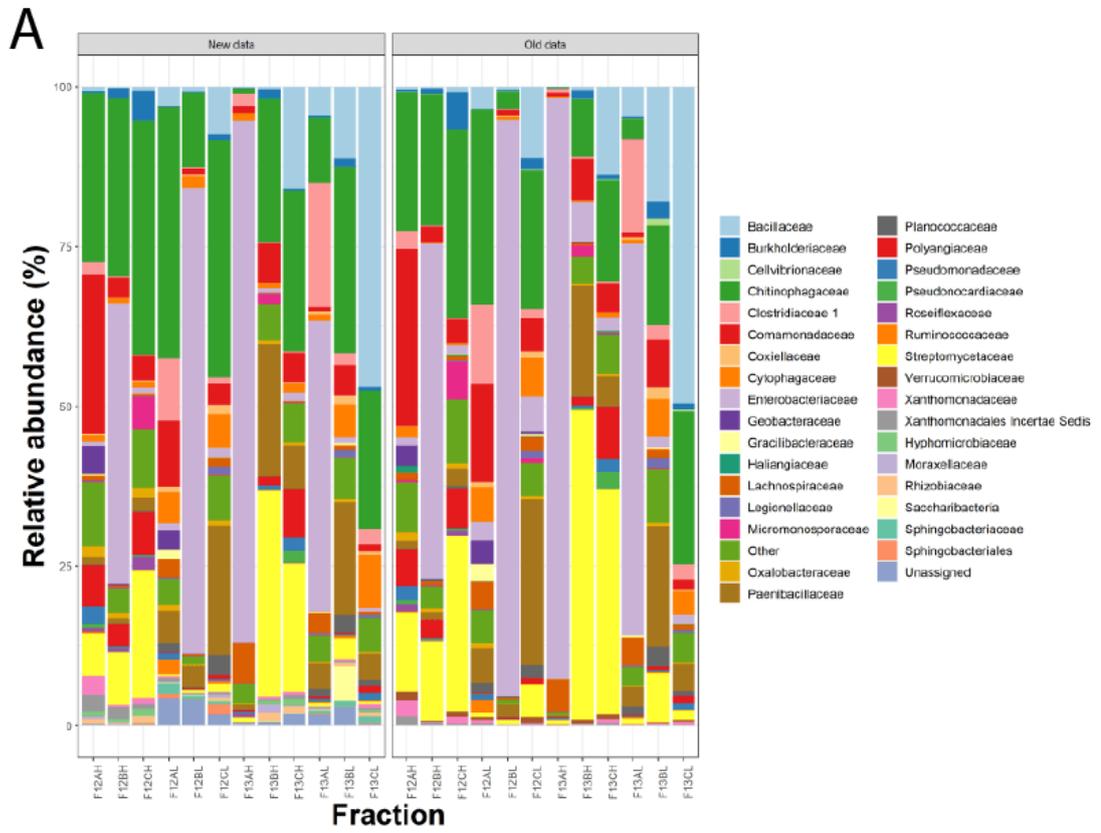


Figure 5.6 Analysis of the SIP 16S rRNA gene sequencing data from the endosphere. **A** Graph demonstrating high variability across SIP fractions from the endosphere. Bars show the relative abundance of each bacterial group within the pooled sequenced  $^{12}\text{C}$  heavy,  $^{12}\text{C}$  light,  $^{13}\text{C}$  heavy and  $^{13}\text{C}$  light fractions (N=3), for two separate sequencing runs on the same samples (old data & new data). All taxa with a median relative abundance < 0.05% were sorted into the “Other” category, and ASVs are presented at the family level. Where family level assignments were unavailable the next taxonomic level is presented. **B** Principal coordinates analysis (PCoA) on Bray-Curtis dissimilarities for the endosphere, performed to assess the consistency of the replicates for endosphere SIP sequencing. Dots show each replicate for  $^{12}\text{C}$  heavy (orange circles),  $^{12}\text{C}$  light (green triangles),  $^{13}\text{C}$  heavy (blue squares), and  $^{13}\text{C}$  light (purple crosses) fractions (N=3).

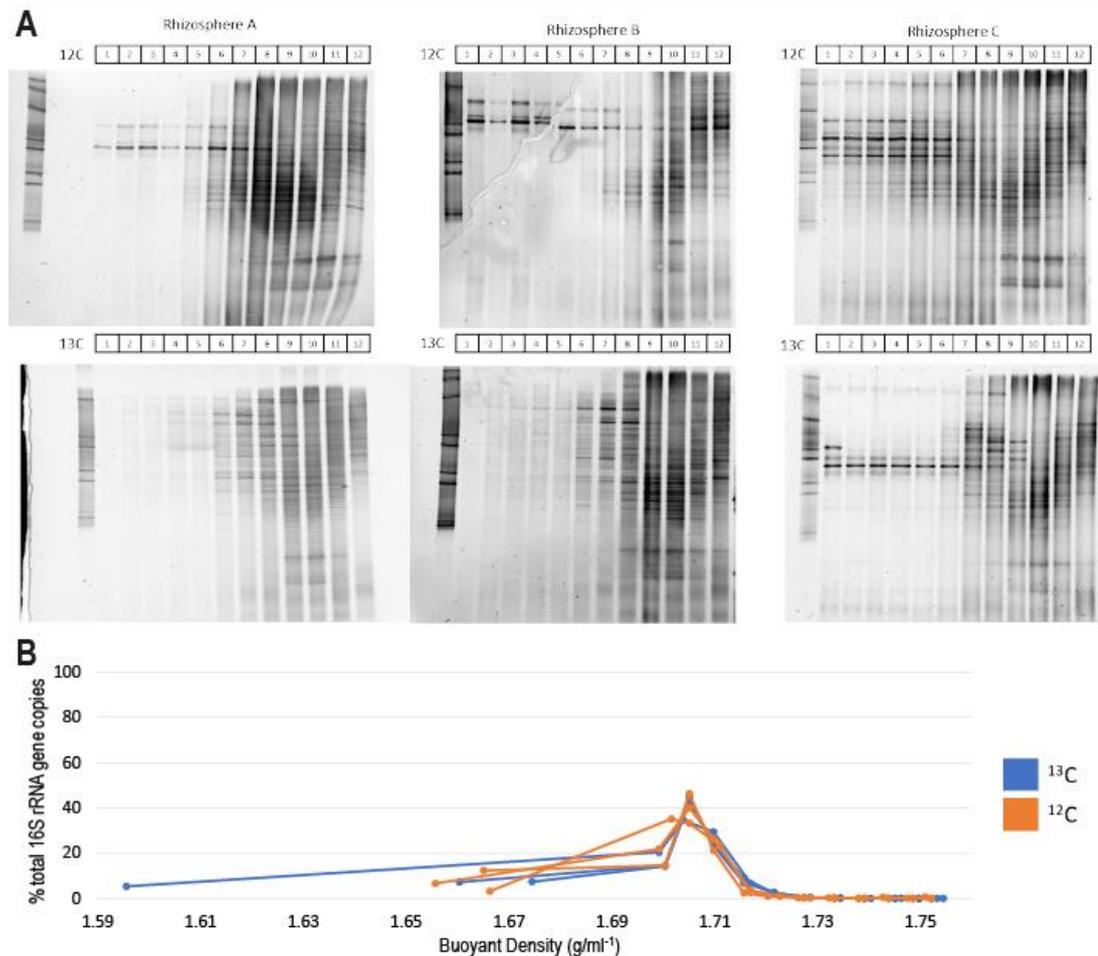


Figure 5.7 Assessment of bacterial 16S rRNA gene sequence diversity and abundance across rhizosphere SIP fractions. **A** Denaturing gel gradient electrophoresis (DGGE) performed to assess bacterial 16S rRNA gene diversity across the 12 fractions generated for the rhizosphere associated with the  $^{12}\text{C}$  control (top) and  $^{13}\text{C}$  labelled (top) plants (N=3). **B** Quantitative PCR against the bacterial 16S rRNA gene, to test for a shift in the abundance of bacteria towards the heavy fractions, which would be indicative of  $^{13}\text{C}$  incorporation. Graph shows the percent of total 16S rRNA genes found within each of the 12 fractions for each rhizosphere sample.  $^{12}\text{C}$  control plants are shown in orange, and  $^{13}\text{C}$  labelled plants are shown in blue (plotted as buoyant densities for that fraction in g / ml<sup>-1</sup>) (N=3).

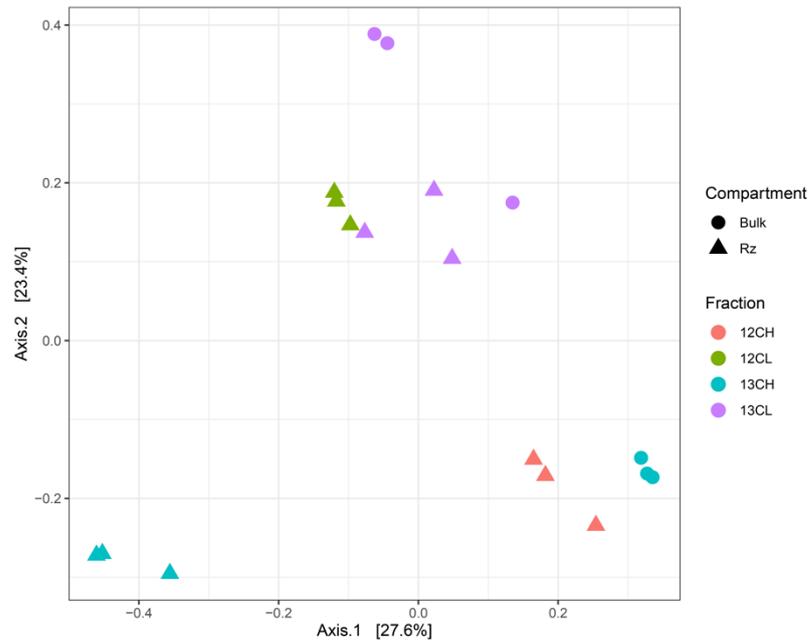


Figure 5.8. Principle coordinates analysis (PCoA) on Bray-Curtis dissimilarity, performed to assess the consistency of the bacterial community across replicates, and to identify a community shift between the different treatments. Points show the  $^{13}\text{C}$  unplanted soil control (circles) or the rhizosphere samples (triangles). For rhizosphere and the bulk soil control, the  $^{13}\text{C}$  heavy (13CH, blue) and  $^{13}\text{C}$  light (13CL, purple) fractions are shown, and for the rhizosphere  $^{12}\text{C}$  heavy (12CH, red) and  $^{12}\text{C}$  light (12CL, green) are shown (N=3 replicate plants for each). Rhizosphere communities were significantly different when comparing between labelled or unlabelled fractions (permanova: permutations=999,  $R^2 = 0.59$ ,  $p < 0.001$ ).

### 5.3.3 Rhizosphere compartment

Both DGGE and qPCR showed that there was a shift in the bacterial community in the rhizosphere towards the heavy fraction, for the  $^{13}\text{C}$  labelled plants (Figure 5.7). PCoA showed that the bacterial community in the  $^{13}\text{C}$  heavy fraction was distinct from that of the control  $^{12}\text{C}$  heavy fraction and was also distinct from both  $^{13}\text{C}$  and  $^{12}\text{C}$  light DNA fractions. This indicated that microorganisms within the community have incorporated the  $^{13}\text{C}$  label. Additionally, the  $^{13}\text{C}$  heavy rhizosphere community was also distinct from the  $^{13}\text{C}$  unplanted bulk soil control (Figure 5.8), strongly indicating that labelling within the rhizosphere was the result of host-derived carbon. These observations were corroborated via permanova analysis, which showed a significant shift in bacterial community composition from the  $^{13}\text{C}$  heavy fraction compared to the  $^{12}\text{C}$  control fractions, (permanova:  $R^2 = 0.59$ ,  $p < 0.001$ ). These comparisons strongly indicated that the shift in community composition within the  $^{13}\text{C}$  heavy DNA fraction was the result of microbes within the rhizosphere actively utilising  $^{13}\text{C}$  labelled root exudates.

Differential abundance analysis was performed to identify the taxa driving the shifts within the  $^{13}\text{C}$  heavy fraction. Exudate metabolisers were defined as taxa showing significantly greater abundance within  $^{13}\text{C}$  heavy DNA fractions when compared with both the  $^{13}\text{C}$  light fractions and the  $^{12}\text{C}$  control heavy fractions. Above the abundance threshold, we identified nine exudate-utilising bacterial taxa (Figure 5.9). While *Streptomycetaceae* were not among these, three other core enriched bacteria were found to utilise root exudates, *Pseudomonadaceae*, and both *Comamonadaceae* and *Oxalobacteriaceae*, which belong to the *Burkholderiaceae*. As defined by the Genome Taxonomy Database <sup>304</sup>, *Comamonadaceae* and *Oxalobacteriaceae* are now classified as genera *Comamonas* and *Oxalobacter* within the *Burkholderiaceae* family.

Six other taxa were also found to utilise root exudates, *Verrucomicrobiaceae*, *Enterobacteriaceae*, *Micrococcaceae*, *Paenibacillaceae*, *Cytophagaceae* and *Fibrobacteraceae*. The most abundant of these taxa were the *Enterobacteriaceae*, though this group was not identified within any of the datasets discussed in chapters three and four, except for the seed endosphere (Chapter Four). The parent class for this family however, the Gammaproteobacteria, was identified in all root samples, but was excluded from differential abundance analysis due to its high-level taxonomic identification. To explore whether these Gammaproteobacteria ASV's could belong to the *Enterobacteriaceae* family, the reads were extracted and ran through NCBI BLAST <sup>292</sup>; this however did not yield any alignments with an identity >95%, and thus revealed no additional information about the identity of the Gammaproteobacteria reads.

While not identified as a core-enriched taxa, in chapters three and four *Micrococcaceae* were detected at low quantities within the roots of all plants cultivated within agricultural soil; whilst this family constituted a small percentage of the microbial community within Levington F2 compost, *Micrococcaceae* were not detected within the endosphere of plants cultivated in Levington F2 compost, indicating that they were only able to colonise the root from agricultural soil (Figure 4.2). In chapter three, *Fibrobacteraceae*, *Cytophagaceae*, and *Paenibacillaceae* were all identified by differential abundance analysis as candidate core enriched endosphere or rhizosphere taxa, as all showed a significant increase in their abundance within the root associated compartments regardless of soil type (Supplementary Table S.3). They were abandoned as candidate core enriched taxa however as their abundance fell below the threshold that was selected to exclude false positives resulting from low abundance taxa. Their identification as exudate

utilisers provides limited evidence that these three taxa may indeed be core enriched members of the root community. Further, *Paenibacillaceae* were enriched within the rhizosphere at the stem elongation growth phase (Figure 3.5 A). The abundance of this group within the endosphere was significantly lower after senescence, the same pattern was observed for core enriched exudate utilisers *Pseudomonadaceae* and *Burkholderiaceae* (Figure 3.5 A). Together this indicates that taxa reliant on root exudates may be unable to persist within the root after developmental senescence.

Described in chapter four are strain to species level identifications of exudate utilising taxa within the endosphere, including for the families *Pseudomonadaceae*, *Burkholderiaceae*, *Micrococcaceae*, *Paenibacillaceae*, and *Cytophagaceae* (Chapter Four, Section 4.5). Within chapter four, one exudate utilising genus was detected within the roots by long-read amplicon sequencing, *Niastella* from the *Chitinophageaceae* family. Surprisingly however this family was not found to utilise root exudates, despite the prevalence of the *Chitinophageaceae* within wheat roots in all previously described data sets, and within the endosphere fractions (Figure 5.6) and within the unfractionated data (Supplementary Figure 5).

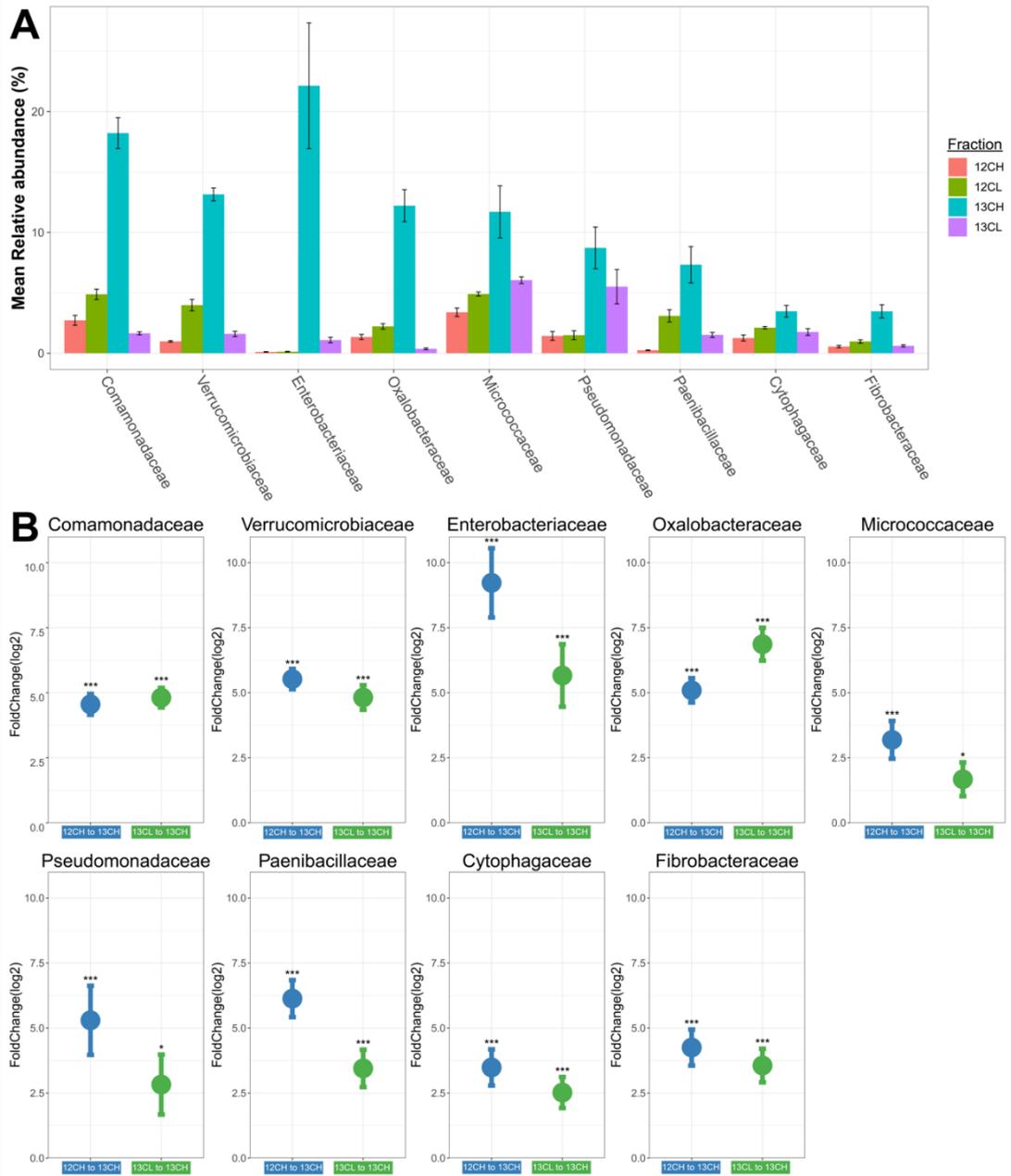


Figure 5.9. Graphs showing the relative abundance of exudate-utilisers across the different samples, and differential abundance analysis demonstrating exudate utilisation. **A** shows mean relative abundance of each root exudate utilising bacterial family in the rhizosphere of plants incubated with  $^{12}\text{CO}_2$  or  $^{13}\text{CO}_2$  (N=3 replicate plants per treatment). Error bars show  $\pm$  standard error of the mean. **B** The results of differential abundance analysis for bacterial families in the rhizosphere; points show the  $\log_2$  fold change of each of the bacterial families between the  $^{12}\text{CO}_2$  heavy and the  $^{13}\text{CO}_2$  heavy fraction (blue) or between the  $^{13}\text{CO}_2$  light and the  $^{13}\text{CO}_2$  heavy fraction (green) (N=3). Bars show  $\log_2$  fold change standard error (N=3), \* indicates  $p < 0.05$ , \*\*\* indicates  $p < 0.001$ , Bonferroni corrected. For the full statistical output see Supplementary Table S.8.

## 5.4 Discussion and conclusions

No exudate utilisation was detected for the archaeal community. Few studies have explored root exudate utilisation by archaea, though Bressan *et al.* reported archaeal exudate utilisation for a mutant line of *Arabidopsis thaliana*<sup>69</sup> which overproduced hydroxybenzylglucosionlate, a plant metabolite stored in the vacuole. This compound is involved in pathogen defence, under stress glucosionlates are hydrolysed into glucose, sulphates, and a range of biocidal compounds. The archaeal community was not generally influenced by hydroxybenzylglucosionlate overproduction, though this may be the result of the archaeal database limitations discussed in chapter 3. The authors do report however that SIP revealed one archaeal group (an uncultured crenarchaeote) had incorporated the <sup>13</sup>C label within the rhizosphere of this mutant line. This could indicate that this group was utilising host metabolites resulting from hydroxybenzylglucosionlate overproduction; equally crenarchaeota are known as CO<sub>2</sub> fixing autotrophs<sup>248</sup>, and so could be labelled autotrophically within this experiment. Were this the case however it would not explain why they were only labelled in rhizosphere of the glucosionlate overproduction mutant.

The best studied example of archaeal exudate utilisation is within the rhizosphere of rice<sup>191</sup>. This is a starkly contrasting environment to the wheat root rhizosphere as rice is cultivated in flooded fields which generates anaerobic conditions within the rhizosphere, and a community dominated by methanogenic archaea. Indeed, root exudate utilisation was detected for RC-1 archaeal methanogens, though this may have been indirect; the authors hypothesised that as exudates were degraded anaerobically by the rhizosphere community H<sub>2</sub> was produced as a bi-product, then used as an energy source by methanogens who are then able to assimilate <sup>13</sup>CO<sub>2</sub> and generate methane. This example then is not evidence for direct exudate utilisation by archaea, rather an indirect effect of root exudation influencing root associated archaea and providing a niche. Overall, neither the literature nor the present work paints a clear picture of the potential for root associated archaea to utilise host derived carbon. It is likely that the extent of archaeal root exudate utilisation is dwarfed by that of bacteria and fungi, not least because faster growing organisms will be more able to monopolise host derived resources, and because bacteria and fungi typically are more dependent on conventional metabolism. Archaea however are more likely to be chemoautotrophs, for example ammonia oxidising archaea or methanotrophs which are dominant in the agricultural soil used

in this experiment (as shown in Chapter 3), and thus have less need for root exudates.

For the fungal community no labelling of the community was detected. This is surprising given the numerous examples from the literature of fungi utilising root exudates from cereal crops or grasses<sup>59,393–396</sup>, including for wheat<sup>66</sup>. It could be that the two-week labelling period chosen for this experiment, or the 2x atmospheric CO<sub>2</sub> concentration chosen for growth chambers, was not sufficient for the fungal community to become labelled; this is unlikely however as all previous studies that detected <sup>13</sup>C labelled fungi within the rhizosphere used shorter labelling periods and/or the same or lower CO<sub>2</sub> concentrations<sup>59,66,393–396</sup>. Another possibility is that within the church farm agricultural soil, the fungal community is not active or does not utilise host-derived carbon, though this seems unlikely given the quantity of fungi detected within the roots of wheat and the influence of compartment of wheat community composition (Chapter 4, Figures 4.1 and 4.4). Wheat is thought to display a relatively weak rhizosphere effect<sup>68</sup>, which could also be a factor, limited exudation rates could make it more challenging to observe exudate utilisation by slower growing fungi when fast growing gram negative bacteria are more able to monopolise resources and more rapidly incorporate the <sup>13</sup>C label. Had we incubated plants in <sup>13</sup>CO<sub>2</sub> for longer to account for this, newly labelled taxa would have been indistinguishable from those labelled as a result of cross-feeding, however. It could also be that the growth of root-associated fungi is supported in other ways, for example the shedding of BLC's could provide carbon to support beneficial fungi within the rhizosphere<sup>321</sup>, as discussed in chapter 3. After two weeks of incubation with <sup>13</sup>CO<sub>2</sub> these BLC's would not have been labelled, as the turnover of these cells takes four weeks, so if this were the case the present experiment would not have detected fungal utilisation of this host-derived carbon.

It is possible that the specific soil system used in this study does not harbour beneficial root associated fungi which are able to utilise root exudates. The likelihood of this hypothesis is difficult to postulate given the high-level taxonomic identifications for fungal community profiles in chapters 3 and 4, but the cultivation methods and soil conditions used may provide some useful clues. For example beneficial arbuscular mycorrhizal fungi associate more strongly with plant roots under extreme abiotic stress, such as drought stress within desert ecosystems<sup>397</sup>, such conditions were not present for laboratory cultivated plants. Fertilisation practices (particularly phosphate treatments) also reduce the benefits that many crops obtain from mycorrhizal fungi<sup>398,399</sup>. In legumes, while arbuscular mycorrhizal

fungi improve the efficiency of nitrogen assimilation by *Rhizobia*, domestic varieties recruit less diverse communities of beneficial fungi to the root, and some are unable to recruit arbuscular mycorrhizal fungi at all <sup>71</sup>. Among some crop species then domestication has negatively impacted the ability of the plants to recruit beneficial fungi, and in some soils, they are unable to recruit any beneficial fungal microbiota at all. Indeed, domestication is also thought to have negatively impacted the benefits that cereal crops can gain from arbuscular mycorrhizal fungi <sup>70</sup>. Additionally there is a positive association between the diversity of arbuscular mycorrhizal fungi and both soil organic carbon and available nitrogen levels <sup>399</sup>; soil chemical analysis showed that the church farm agricultural soil used in the present work had relatively low levels of inorganic nitrogen, and very low levels of organic matter (Chapter 2, Table 2.3). This could indicate that the agricultural soil does not provide a favourable environment for mycorrhizal fungi. Multiple factors that are known to negatively impact the colonisation efficacy and activity of arbuscular mycorrhizal fungi are at play within agricultural systems, and the system used for this experiment. In the present work plants were cultivated under laboratory conditions, and thus were not subjected to any abiotic stress. The soil used was poor in nutrient and organic matter and was freshly sampled from a commercially active agricultural site, and the experiment also used a domesticated variety of wheat. Cumulatively these factors may have all influenced the diversity and activity of mycorrhizal fungi available to the host plant and may explain why no fungi were labelled in the present experiment.

Variability within the endosphere was too high for the present work to draw any conclusions about the utilisation of host derived carbon within the root interior. This is especially unfortunate as chapters 3 and 4 show that the endosphere microbial community composition for bacteria and fungi was significantly different to that of the rhizosphere, particularly due to the colonisation of *Streptomycetaceae*. It was hypothesised that a different subset of microbes to the rhizosphere, and particularly that *Streptomycetaceae*, would be utilising host metabolites within the endosphere, however the present work can contribute no evidence to support this hypothesis. This variability is surprising given the consistency of the endosphere bacterial community demonstrated within other experiments using this soil type, for example across the different varieties of wheat shown in chapter 4. Unlike previous experiments using this soil type, there was a high relative abundance for the *Enterobacteriaceae* within these samples, a family not previously observed within the endosphere by this work within this soil type or for this variety of wheat. It was

hypothesised that this may be the result of sample contamination at the sequencing facility; the samples for the endosphere were therefore re-sequenced, however this returned the same results (Figure 5.6). The bacterial 16S rRNA gene community diversity was assayed for these same plants but using unfractionated DNA. This showed that the *Enterobacteriaceae* were prevalent to a high degree within the endosphere for just one replicate (Supplementary Figure S.4). This anomalous result is likely to be responsible for the high variability observed within the endosphere, and thus is likely to be the reason that the endosphere data was uninterpretable. Further to this, our ability to consistently measure the microbial community composition within the endosphere is likely to be particularly sensitive to the fractionation process when compared to the rhizosphere; the bacterial load is significantly lower within the endosphere compared to the rhizosphere, and so dividing this by twelve fractions during the fractionation process may increase the stochastic effects on community composition, thus reducing the consistency of biological replicates. Is it clear however from the unfractionated data that the inconsistent colonisation of *Enterobacteriaceae* is likely to be responsible for the high variability within the endosphere data. In the future more replicates could be sampled to account for this variability within the endosphere, or a larger quantity of DNA per sample could be used for density gradient ultracentrifugation and fractionation.

Nine bacterial taxa were found to utilise root exudates, including core enriched taxa such as *Burkholderiaceae* (*Comamonadaceae* and *Oxalobacteriaceae*) and *Pseudomonadaceae*, this observation supports previous studies<sup>65,197</sup>. As these bacterial groups were consistently enriched within the rhizosphere or endosphere across all soil types and varieties of wheat tested, these results strongly imply these families may be selectively recruited to the plants via root exudates. The *Pseudomonadaceae* family contains a diverse range of plant-beneficial and plant pathogenic strains<sup>400,401</sup>, the literature however correlates exudate utilisation with microbial functions that benefit the host plant<sup>142,343</sup>, and root exudates can have a negative effect on plant pathogens<sup>402</sup>. While the mechanism of this selectivity remains ambiguous, it is likely that these exudate utilisers are plant beneficial strains. For *Pseudomonadaceae* there are well studied examples with plant growth promoting traits such as *Pseudomonas brassicacearum*<sup>403</sup> and *Pseudomonas fluorescens*<sup>404</sup>. *Pseudomonas poae* is also a plant beneficial species from this family<sup>372</sup>, and the full length 16S rRNA gene sequencing presented in chapter four identified this species within the endosphere and rhizosphere (Chapter 4.5, Figure

4.13). Thus, it is likely that *P. poae* is the pseudomonad utilising host derived carbon within the rhizosphere.

Within the *Burkholderiaceae*, *Comamonas testosteroni* has been shown to increase nutrient bioavailability for wheat via nitrogen fixation and potassium solubilisation <sup>405</sup>, and in conjunction with silver nanoparticles can also help alleviate salt stress in Flax (*Linum usitatissimum*, cultivated for linseed) <sup>30</sup>. No previous association between *Oxalobacter* and plants has been reported. Interestingly these bacteria are well known oxalotrophs <sup>406</sup>; oxalic acid has broad biological functions and is also a common root exudate compound, and oxalotrophy is a common trait of plant-beneficial *Burkholderia* <sup>407</sup>. It is possible that the labelling of *Oxalobacter* is the result of oxalotrophy within the rhizosphere. As has been mentioned however, the *Oxalobacteraceae* and *Comamonadaceae* are now categorised as genera within the *Burkholderiaceae* family, and the genera formerly contained within these families have also been recategorized into the *Burkholderiaceae* family (GTDB <sup>304</sup>). Long-read amplicon sequencing presented in chapter four identified a number of genera formerly categorised as *Oxalobacteraceae* (*Duganella*, *Massilia*, and *Herbaspirillum*) or *Comamonadaceae* (*Rhodoferax* and *Polaromas*), and *Duganella* and *Massilia* were the most abundant genera from these groups within the endosphere (Chapter Four, Section 4.5, Figure 4.11). the *Burkholderiaceae* family genera *Massilia*, *Rhodoferax*, and *Variovorax* were all detected within the rhizosphere. Given that it was *Oxalobacteraceae* and *Comamonadaceae* were identified as exudate utilising families by this experiment, it is likely that *Massilia* and *Rhodoferax* are responsible for exudate utilisation from these two groups respectively.

Six other taxa, which had not been identified as core root-associated taxa, were also found to utilise root exudates, *Cytophagaceae*, *Enterobacteriaceae*, *Fibrobacteraceae*, *Micrococcaceae*, *Paenibacillaceae*, and *Verrucomicrobiaceae*. A number of these have, to different degrees, demonstrated an ability to provide benefits to the host plant and thus are likely to represent beneficial microbiota. For example, within the *Micrococcaceae*, the genus *Arthrobacter* can support plant growth <sup>408–410</sup> in a number of ways including by alleviating salt or desiccation stress <sup>28</sup>, and has been shown to colonise wheat tissues in laboratory experiments <sup>411</sup>. As shown by the full length 16S rRNA gene sequencing presented in chapter four (Chapter 4, Section 4.5, Figure 4.16), this genus was present within the endosphere of wheat, and so it is likely that *Arthrobacter* were utilising root exudates within this experiment. Within the *Verrucomicrobiaceae*, the genera *Roseimicrobium* and

*Verrucomicrobium* have both been isolated from the rhizosphere <sup>412,413</sup>, and *Verrucomicrobium* can help relieve cadmium stress as a part of a microbial consortium <sup>57</sup>. No genera from this family were detected within the roots by long-read sequencing.

Concurrent with the present findings, *Paenibacillaceae* were also identified as root exudate utilisers within the wheat root microbiome by Uksa and colleagues <sup>196</sup>. This family contains seven genera (GTDB <sup>304</sup>), including plant associated genera such as *Cohnella*, which have been isolated from the root nodules of runner bean <sup>414</sup> (*Phaseolus coccineus*), and the Mediterranean pulse white lupin <sup>415</sup> (*Lupinus albus*). Other plant associated *Paenibacillaceae* family genera include *Saccharibacillus* <sup>212,416,417</sup>, which has also been isolated from the wheat endosphere <sup>418</sup>, and can produce cellulolytic enzymes <sup>419</sup>, and while not typically plant associated, the genus *Thermobacillus* is known for its ability to degrade plant cell walls <sup>420</sup>. The most well-known genus within this family is *Paenabacillus*, species within this genus are well known for their plant growth promoting activity such as nitrogen fixation <sup>421</sup> and pathogen suppression <sup>422,423</sup>. The long-read amplicon sequencing presented in chapter 4 however (Chapter 4, Section 4.5, Figure 4.16) showed that *Cohnella* were present within the rhizosphere of wheat, and so this is likely to be the *Paenibacillaceae* genus utilising root exudates.

Within the *Cytophagaceae* are a number of plant-beneficial species, for example *Pontibacter niistensis*, which can promote the growth of a range of plants. Genera such as *Dyadobacter* have been isolated from plant tissues <sup>424</sup>, and the species *Cytophaga hutchinsonii* can degrade crystalline cellulose <sup>425</sup>. This species was also identified within the endosphere by the long-read sequencing experiment presented in chapter four, and so is likely to be the species responsible for exudate utilisation by the *Cytophagaceae* (Chapter 4, Section 4.5, Figure 4.16). This long read 16S rRNA gene sequencing did not identify any specific *Fibrobacteraceae* genera within the root. Within the family *Fibrobacteraceae*, *Fibrobacter* are typically associated with cellulose degradation <sup>426</sup>, and so could occupy a plant associated niche.

Overall, for the majority of exudate-utilising families, *Comamonas*, *Pseudomonadaceae*, *Micrococcaceae*, *Verrucomicrobiaceae*, *Paenibacillaceae*, and *Cytophagaceae*, there are examples within the literature of plant-beneficial species for these families, thus it is likely that root exudates selectively recruit and support the growth of these beneficial taxa within the rhizosphere. For some groups however there is currently no evidence for plant-beneficial functions within the

literature, for example for *Fibrobacteraceae*. Cellulose or plant cell wall degradation however is a common trait within the exudate utilising families. This could indicate several things. Firstly, it could be that within this experiment labelling for some taxa was the result of cell wall consumption and not root exudate utilisation. Given the two-week labelling period however it is unlikely that plant cell wall components would have been labelled to a sufficient degree for cellulose, hemicellulose, or lignin consumption from the root to result in labelling. When trying to label wheat root tissues with  $^{13}\text{CO}_2$  Bernard and colleagues incubated plants with 400 ppmv  $^{13}\text{CO}_2$  for 35 days <sup>187</sup>, further supporting that it is unlikely that tissues would have been labelled to a sufficient extent during a 14 day labelling period. Alternately, cellulose degradation could simply be a common trait amongst successful and plant-beneficial rhizosphere bacteria. Given the large quantities of discarded plant matter associated with wheat roots, resulting from shedding of BLC's <sup>321</sup>, self-destructive defensive mechanisms such as NETs <sup>322</sup>, and frequent turnover of lateral root cells <sup>323</sup>, it stands to reason that microbes within this environment would be capable of utilising this rich energy source. Indeed, it could be that plant roots in part use this as a mechanism to select for beneficial root microbiota.

The most abundant exudate utilising family, *Enterobacteriaceae*, contains many well-known human pathogenic genera such as *Escherichia*, *Klebsiella*, *Salmonella*, and *Yersinia*, and are not typically studied in a plants-associated context. It was postulated that the presence of this family was a contaminant, given the anomalous results from the endosphere. However, unlike the endosphere, within the unfractionated rhizosphere samples the abundance of *Enterobacteriaceae* consistent across replicates (Supplementary Figure S.4). Thus, it must be concluded that the presence of *Enterobacteriaceae* within these samples is an honest reflection of the rhizosphere community for the plants used for this experiment. It is not clear which *Enterobacteriaceae* taxa are likely to be utilising root exudates, it is possible that these are novel plant associated *Enterobacteriaceae* strains, previously unstudied.

Despite the prevalence of *Streptomycetaceae* within the endosphere, DNA-SIP failed to confirm that this family is able to utilise root exudates within the rhizosphere. While this is concurrent with observations in *Arabidopsis* by Worsley and colleagues <sup>137</sup>, this contradicts the findings of Ai and colleagues <sup>65</sup>. Most of the exudate utilising families identified in the present work were fast growing Gram-negative bacteria. As observed by Worsley and colleagues <sup>137</sup>, faster growing organisms are labelled more readily within a two-week incubation period. Due to

their faster growth rates, these microorganisms can more easily monopolise the plant derived carbon within the rhizosphere and incorporate  $^{13}\text{C}$  into the DNA backbone during DNA replication. It is possible that slower growing organisms such as *Streptomycetaceae* are outcompeted for root derived resources in the rhizosphere by faster growing gram-negative organisms. It must also be noted that the good quality data for the DNA-SIP experiment came from the rhizosphere, *Streptomycetaceae* however primarily colonised the endosphere and thus may primarily consume host-derived carbon from the root interior.

Core to the design of this experiment is the assumption that actively growing microorganisms replicate their DNA and thus incorporate  $^{13}\text{C}$  into the DNA backbone. Actively growing streptomycetes however, similarly to many filamentous fungi, do not replicate their DNA as often as unicellular bacteria because *Streptomyces* species and filamentous fungi grow primarily through hyphal tip extension<sup>427</sup>. While unicellular bacteria replicate their chromosome at each cell cycle during binary fission, the typical mode of growth for *Streptomyces* spp. is hyphal tip extension, which progresses via the elongation of vegetative cells which contain multiple copies of the chromosome. These cells then septate, and chromosomal distribution can be stochastic. Chromosomal replication in *Streptomyces* spp. is most rapid during aerial hyphae formation & sporulation, a heavily regulated processes later in the life cycle<sup>428</sup>. Fungi share the same stochastic chromosomal distribution, as they grow via hyphal tip extension within multinuclear hyphae, so chromosomal replication is likely to also be less frequent than for unicellular bacteria<sup>429</sup>. This indicates that typical *Streptomyces* or fungal growth mechanisms may mean that exudate utilisation is difficult to detect via DNA SIP, as chromosomal replication is not occurring to the same extent as for unicellular bacteria. In future experiments then RNA SIP should be used to account for this, as RNA turnover is more rapid, and all active microorganisms, regardless of the mode of growth, will more easily become labelled with  $^{13}\text{C}$  from host derived metabolites. This approach may resolve the limitations both for filamentous bacteria like *Streptomyces*, and also for the fungal community in future SIP studies.

Further SIP experiments exploring the endosphere community, with more replicates to account for the high variability within the endosphere, may help to determine whether *Streptomycetaceae* can utilise plant derived carbon, and if the loss of these resources explains their reduced relative abundance within the endosphere during senescence. As has been discussed for fungi and other groups, *Streptomycetaceae* could be utilising host-derived carbon that was not labelled during the two-week

labelling period used in this experiment. *Streptomyces* are known to degrade hemicellulose<sup>326</sup>, and as discussed the root environment is rich in carbon from discarded plant matter resulting from programmed cell death in lateral root cells, BLC shedding, and root border cell death caused by NET deployment.

During developmental senescence, nitrogen is the main resource diverted to the developing grain<sup>223</sup>, and during this final stage of the plants life *Streptomycetaceae* endosphere abundance decreased over two-fold (Chapter 3). It is possible then that nitrogen, not carbon, is the resource provided by the host plant to support *Streptomycetaceae* growth within the endosphere or rhizosphere. There is precedent for host-derived metabolites such as amino acids or gamma-aminobutyric acid (GABA) acting as a nitrogen source for root associated microbes<sup>141,142</sup>. Additionally, the increased use of nitrogen fertilizer (which correlates with greater total root exudation) was negatively correlated with the abundance of *Streptomycetaceae* in the wheat rhizosphere<sup>96</sup>, implying the possibility of reduced reliance on host derived nitrogen after nitrogen fertiliser is added to the soil. In the future, <sup>15</sup>N-nitrogen DNA or RNA-SIP could be used to explore whether *T. aestivum* var. Paragon is able to support *Streptomycetaceae*, or any other taxa within the endosphere, via nitrogen containing, host-derived metabolites.

In summary, this chapter has identified nine exudate utilising bacteria within the rhizosphere of wheat, *Pseudomonadaceae*, *Verrucomicrobiaceae*, *Enterobacteriaceae*, *Micrococcaceae*, *Paenibacillaceae*, *Cytophagaceae* and *Fibrobacteraceae*, and from the *Burkholderiaceae* family, *Oxalobacteraceae* and *Comoamonadaceae*. It is likely that these taxa are selectively recruited to the rhizosphere and maintained via root exudates, and that they provide some benefit to the host. Indeed, reviewing the literature for these groups revealed that the majority of these taxa contain bacterial species that can provide benefits to plants such as abiotic stress relief, pathogen defence, and increased nutrient bioavailability. For *Enterobacteraceae* no precedent in the literature could be found for plant-interactions. No exudate utilisation was detected from the archaeal or the fungal community, and data from the endosphere was too variable for any conclusions can be drawn about bacterial utilisation of host derived carbon in the endosphere. Surprisingly exudate utilisation was not detected for *Streptomycetaceae*, it could be that these bacteria utilise a different carbon source within the root such as hemicellulose, or that plants provide nitrogen to support *Streptomycetaceae* within the root, and not carbon. Overall, this chapter has provided a solid basis for future studies investigating the role of these nine bacterial taxa within wheat root

microbiomes and raises important questions about the maintenance of *Streptomyetaceae* within wheat roots.

## Chapter 6. Isolation and characterisation of *Burkholderiaceae* and *Pseudomonas* endophytes

### Aims

16S rRNA gene metabarcoding indicated that, after *Streptomyces*, *Burkholderiaceae* were among the most abundant core endosphere taxa associated with the endosphere of wheat. The 300bp V3-V4 amplicon used for sequencing however is insufficient to identify these endophytes beyond the family level; furthermore the *Burkholderiaceae* family contains a wide variety of genera including *Burkholderia*, *Variovorax*, *Rhizobacter*, and *Oxolabacter* (GTDB<sup>430</sup>). This makes the function and identity of these endophytes difficult to postulate based on this data. Prior to the long-read amplicon sequencing described in chapter four this chapter aimed to address two primary questions. (1) Which *Burkholderiaceae* genera are present within the roots of wheat? (2) Can we identify plant beneficial traits for these strains? (3) Can we isolate and identify any other core endosphere taxa such as *Pseudomonadaceae* or *Chitinophagaceae*? To address these questions a combination of targeted isolation, and bioactivity assays were used. Once full length 16S rRNA gene sequences were from the endosphere however the majority of these isolates were concluded not to be prevalent within the endosphere. The data however still demonstrates some interesting capabilities for gram negative endosphere isolates, and thus is included as a supplementary chapter.

### Results

#### 6.1 Isolation and identification of root endophytes

To selectively isolate *Burkholderiaceae* and *Chitinophagaceae* family endophytes, an initial experiment used three media previously used to isolate either *Burkholderia* (BAz<sup>431</sup>, R2A<sup>432-434</sup>) or *Chitinophaga* (MAG<sup>214</sup>). R2A (or Reasoner's 2A agar) is a common rich medium used to isolate *Burkholderia* from both plant and insect associated microbiomes<sup>432,434</sup>, and contains a mixture of simple and complex carbon sources. BAz is a nitrogen-free medium containing azelaic acid as a sole carbon source. This medium was chosen to potentially select for diazotrophic *Burkholderia*, such as have been isolated from other grass species<sup>431</sup>. Many root-associated *Burkholderia* species are capable of oxalotrophy, or the breakdown of the root exudate compound oxalic acid<sup>407</sup>. Thus, it was reasoned that adding oxalic acid to the medium may help select for root associated *Burkholderia* and so variants

of R2A and BAC supplemented with oxalic acid were also used. MAG was chosen as a second rich medium which had previously been used to isolate *Chitinophaga*<sup>214</sup>. While there were no visible colonies on either variant of the BAZ medium after ~24 hours incubation at 30°C, in total 308 colonies displaying different morphologies were isolated from R2A (both variants) and MAG. Of these 115 were screened using an end-point PCR test targeting the *Burkholderia* 16S rRNA gene<sup>269</sup>, 27 of which yielded a positive band, indicating those isolates might belong to the *Burkholderia* genus. Sequencing of the 16S rRNA gene however revealed that only one of the 27 sequenced isolated belonged to the Burkholderiaceae family (genus *Massilia*), and none belonged to the *Burkholderia* genus. Six isolates were from the genus *Delftia* and of the remaining 21, ten belonged to the genus *Bacillus*, and no sequencing data was acquired for the remaining 11. Given that the majority of isolates which yielded a band using these *Burkholderia* specific primers belonged to the *Bacillus* genus, it was concluded that the PCR assay used was not capable of distinguishing *Burkholderiaceae* family from other taxa and thus was not used for further experiments. 16S rRNA gene sequencing was then performed on 11 additional isolates, covering the 11 morphologies observed among the isolates. This identified five additional *Burkholderiaceae* isolates (genus *Achromobacter*), one *Delftia* isolate, and two *Pseudomonas* isolates.

After the first isolation attempt it was unclear whether the culturable diversity of *Burkholderiaceae* endophytes had been saturated. Thus, a second isolation was attempted using different media. R2A medium was used, supplemented with vancomycin to prevent cultivation of gram-positive microbiota such as *Bacillus*. Additionally, BCSA was used as a selective medium for *Burkholderia*, as this medium is used clinically for the diagnostic cultivation of *Burkholderia cepacia*. As a non-selective medium, nutrient agar was also used. Across the three media, an additional 11 *Delftia* and three *Pseudomonas* isolates were acquired, and also four *Stenotrophomonas* isolates. Two additional *Burkholderiaceae* genera were also acquired, one from the *Ralstonia* genus and four from the *Variovorax*.

In total, 16S rRNA gene sequencing showed that 11 isolates were from the *Burkholderiaceae* family (*Achromobacter*, *Massilia*, *Ralstonia*, and *Variovorax*), though none belonged to the genus *Burkholderia*. While no *Chitinophagaceae* isolates were identified, other taxa of interest that were cultivated include *Pseudomonas* (a core enriched rhizosphere taxon) and *Stenotrophomonas* from the family *Xanthobacteriaceae*. Interestingly, the genus *Delftia*, for which this experiment acquired 18 isolates, according to the genome taxonomy database<sup>430</sup> is

now classified as *Comamonas*, within the *Burkholderiaceae* family, a group which was utilising root exudates in the rhizosphere. For a full list of the isolates see Table 7.1, and for media recipes see Chapter Two, Table 2.1.

## 6.2 Bioactivity of isolates

### 6.2.1 Inhibition of the wheat take-all fungus by *Pseudomonas* isolates

Full length 16S rRNA gene sequencing indicated that the *Pseudomonas* isolates MNA132 and RNL304A belonged to the species *Pseudomonas brassicacearum*, a species attributed with being responsible for take-all decline in agricultural soils <sup>40</sup>. To investigate if these strains, and the three other *Pseudomonas* isolates, were able to inhibit the take-all fungus (*Gaeumannomyces graminis var. tritici*) a bioassay was ran to investigate the ability of these strains to inhibit the wheat take-all fungus *Gaeumannomyces graminis var. tritici*. Figure 6.1 shows that all five *Pseudomonas* isolates (MNA132, RNL304, RNL309, RNL311 and RR104) were able to inhibit *G. graminis var. tritici*. A tree presented in chapter four (Chapter Four, Section 4.5, Figure 4.13) showed that all of these pseudomonad isolate 16S rRNA gene sequences clustered with the *P. brassicacaerum* species, indicating they are all closely related to this species of *Pseudomonas*.

**Table 6.1. Isolate strain list**

Full list of the 44 strains isolated from the wheat root endosphere, as identified by 16S rRNA gene sequencing.

Strain No.	Top 16S BLAST hit	Isolation medium
MNA101	<i>Achromobacter</i> (multiple strains & species)	MAG
MNA102	<i>Achromobacter</i> (multiple strains & species)	MAG
MNA103	<i>Achromobacter</i> (multiple strains & species)	MAG
MNA108	<i>Achromobacter</i> (multiple strains & species)	MAG
MNA119A	<i>Achromobacter</i> (multiple strains & species)	MAG
MNA151	<i>Bacillus</i> (multiple strains & species)	MAG
MNA104	<i>Bacillus</i> (multiple strains & species)	MAG
MNA229	<i>Bacillus</i> (multiple strains & species)	MAG
RNA104	<i>Bacillus</i> (multiple strains & species)	R2A
RNA109	<i>Bacillus</i> (multiple strains & species)	R2A
RNA202	<i>Bacillus</i> (multiple strains & species)	R2A
RNA211	<i>Bacillus</i> (multiple strains & species)	R2A
RNA213	<i>Bacillus</i> (multiple strains & species)	R2A
RNA226	<i>Bacillus</i> (multiple strains & species)	R2A
RNA230	<i>Bacillus</i> (multiple strains & species)	R2A

RNL305	<i>Bacillus</i> (multiple strains & species)	R2A
RNL307	<i>Bacillus</i> (multiple strains & species)	R2A
RNL310	<i>Bacillus</i> (multiple strains & species)	R2A
MNL105	<i>Bacillus mycoides</i> (multiple strains)	MAG
BR217	<i>Delftia / Comamonas</i>	BCSA
BR224	<i>Delftia / Comamonas</i>	BCSA
RNA112	<i>Delftia / Comamonas</i>	R2A
RNA119A	<i>Delftia / Comamonas</i>	R2A
RNA119C	<i>Delftia / Comamonas</i>	R2A
RNA201	<i>Delftia / Comamonas</i>	R2A
RNA205	<i>Delftia / Comamonas</i>	R2A
RNA221	<i>Delftia / Comamonas</i>	R2A
RNL101A	<i>Delftia / Comamonas</i>	R2A
RNL102	<i>Delftia / Comamonas</i>	R2A
RNL218A	<i>Delftia / Comamonas</i>	R2A
RNL218B	<i>Delftia / Comamonas</i>	R2A
RNA126	<i>Massilia</i> (multiple strains)	R2A
RNL309	<i>Pseudomonas</i>	R2A
RNL311	<i>Pseudomonas</i>	R2A
RR104	<i>Pseudomonas</i>	R2A
MNA132	<i>Pseudomonas brassicacearum</i> (multiple strains)	MAG
RNL304A	<i>Pseudomonas brassicacearum</i> strain LBUM300	R2A
RR105	<i>Ralstonia</i>	R2A
RR101	<i>Stenotrophomonas</i>	R2A
RR302	<i>Variovorax</i>	R2A
RR304	<i>Variovorax</i>	R2A
RR307	<i>Variovorax</i>	R2A
RR205	<i>Variovorax / Acidovorax</i>	R2A

### 6.2.2 General bioactivity of *Burkholderiaceae* and *Pseudomonas* isolates

To identify isolates with further antimicrobial potential an initial screen tested the ability of each strain to inhibit gram positive, gram negative, or fungal indicator strains. Four strains, RNL304A, MNA132, MNA119A, and RR104 showed activity against the gram-positive indicator *Bacillus subtilis* JH642<sup>435</sup> (Figure 7.2). Three of these strains were *Pseudomonas* (RNL304A, MNA135, and RR104), while MNA119A was a *Comamonas* strain. The three *Pseudomonas* strains showed a large zone of inhibition surrounding the colony, indicative of bactericidal activity. The *Comamonas* strain MNA119A however showed only a very slim zone of inhibition, indicating either bacteriostatic activity against *Bacillus subtilis* JH642, or indirect inhibition as seen through, for example, iron sequestration<sup>436</sup>.

### 6.3 Testing strains for interaction with *Streptomyces*

Given the prevalence of *Streptomycetaceae*, the potential for interactions between root isolates and a *Streptomyces* strain isolated from wheat roots (*Streptomyces* strain CRS3). After an initial screen, an interaction was confirmed for eight isolates on rich medium (NA) and for one strain (MNA119A) on minimal medium (MM) (Figure 7.3). On NA three of the five *Achromobacter* strains (MNA101, MNA103, and MNA108) showed an inability to grow towards *Streptomyces* CRS3; in isolation *Achromobacter* isolates all formed round colonies, whereas when *Streptomyces* CRS3 was present the side of the colony which faces CRS3 formed a flatter edge, indicating the strain was unwilling to grow towards the streptomycete. This was most pronounced for MNA101 and MNA103. On MM a similar pattern was observed for MNA119A, though not as pronounced.

Four out of 12 *Comamonas* appeared to show some interaction with CRS3 in the initial screen. RNA112 and RNA205 all showed a similar inability to grow towards *Streptomyces* CRS3. Interestingly, when grown in isolation RNA205 formed a small colony with a thin film that covered the entire plate, similarly to BR224 and RNL218B. When spotted adjacent to *Streptomyces* CRS3 however the colony morphology changed drastically and was more similar to RNA112. RNL218B and BR224 both grew as a thin film covering the entire plate. Initially it was thought that these strains were growing towards CRS3, however upon comparison with the isolated plate, it seems this is not the case as colony morphology is similar irrespective of the presence of *Streptomyces* CRS3.

The *Ralstonia* isolate (RR105) also showed a similar inability to grow towards *Streptomyces* CRS3, this genus generally known for the plant-pathogenic species *Ralstonia solanacearum*<sup>437</sup>, a soil-borne root infecting bacterium, demonstrating that *Streptomyces* root isolates have an ability to inhibit bacterial plant pathogens.

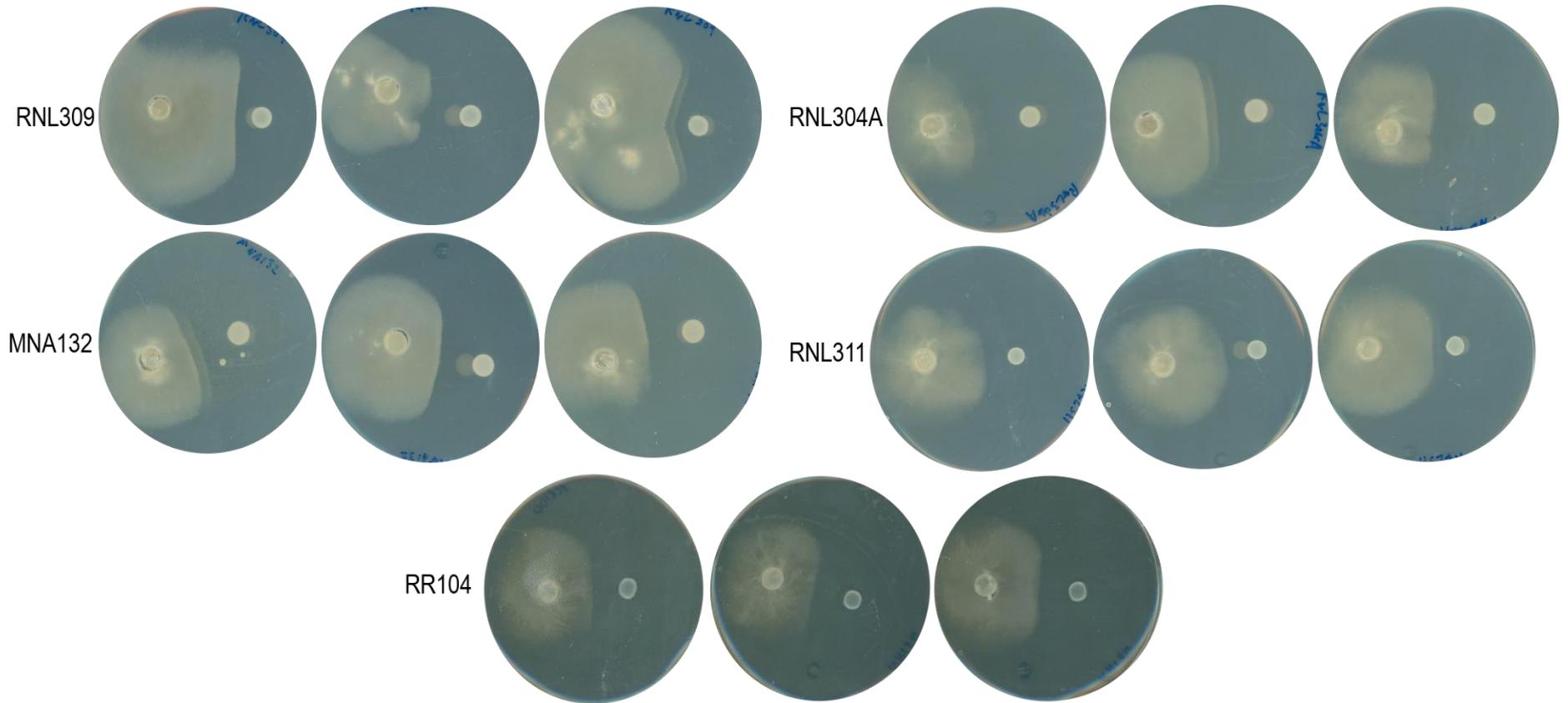


Figure 6.1 Bioassay to test the ability of five *Pseudomonas* isolates to inhibit the wheat take all fungus. Plates show *Pseudomonas* genus root endophyte strains RNL309, RNL304A, MNA132, RNL311 and RR104 demonstrating inhibition of the wheat take-all fungus *Gaemannomyces graminis var. tritici*.

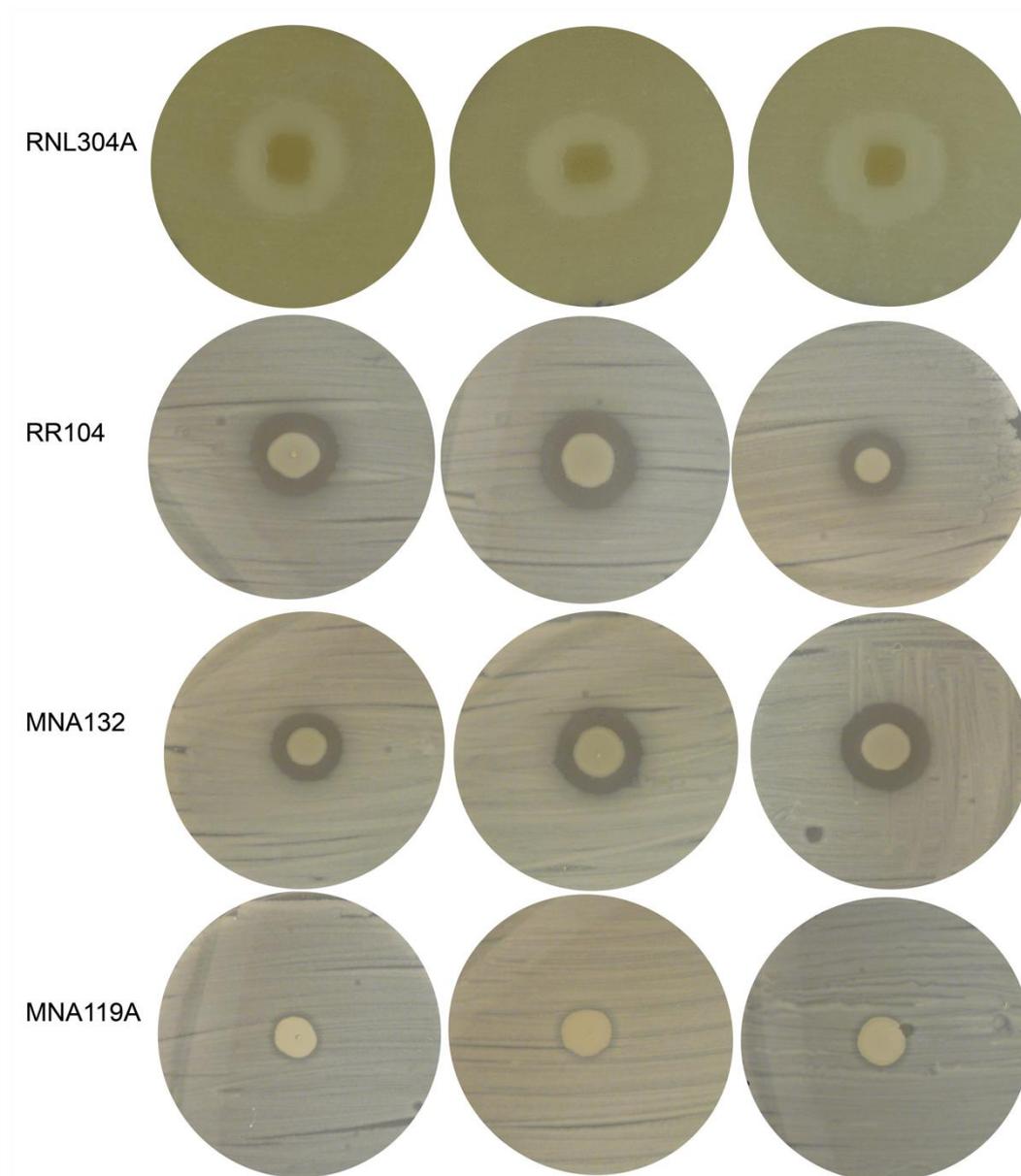


Figure 6.2 Bioassay to test for antibacterial activity from four isolates. Plates show root endophyte strains RNL304A, RR104, MNA132, and MNA119A demonstrating bioactivity against gram-positive indicator organism *Bacillus subtilis* JH642.

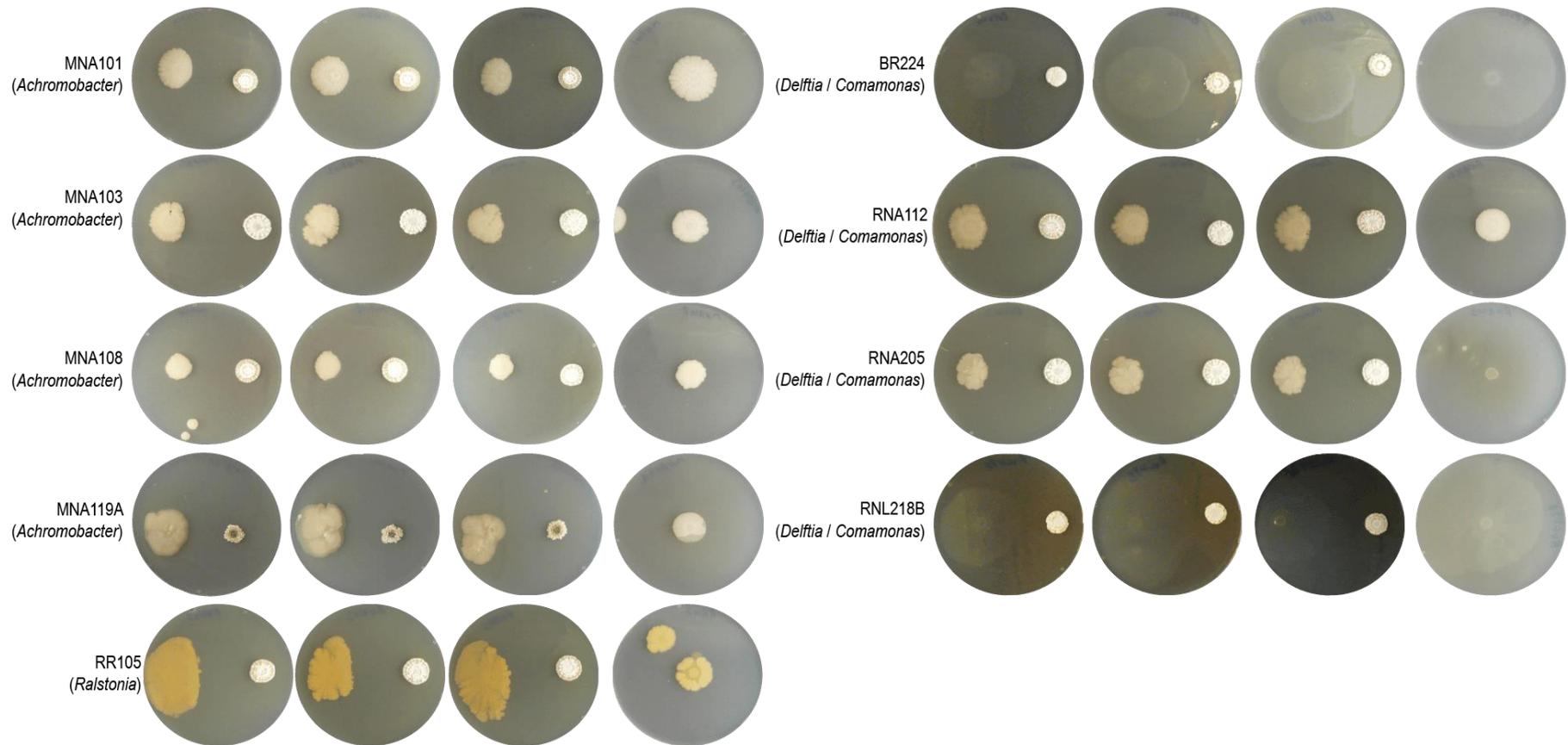


Figure 6.3 Assay to test the ability of endosphere isolates to interact with a *Streptomyces* isolate (CRS3). Plates show the endosphere isolate on the left, and the streptomycete on the right. N=3 per strain, and the fourth plate in each row shows the strain cultivated in isolation, without the presence of *Streptomyces*.

## 6.4 Assessment of phytase activity from isolates

Phosphorus is a key element for plant growth, but many agricultural soils are limited in bioavailable phosphate as the dominant form of phosphate in soils bound to phytate, a highly stable and difficult to degrade molecule<sup>13</sup>. Microbial degradation is required to release bioavailable phosphate into the soil to support plant growth. Many plant-beneficial soil microorganisms are able solubilise phosphate for this purpose using phytase enzymes to breakdown phytate<sup>32</sup>. To investigate if the isolates described here are capable of this activity phytase activity was assessed under standard culture conditions by Gregory Rix according to the methods presented in Rix *et al.*<sup>438</sup>. This did not show any activity from any of the isolates listed in Table 7.1, showing that under laboratory conditions these isolates will not degrade phytate. This could either be because in standard rich medium these isolates do not express any genes related to phytate degradation (such as those encoding for phytase enzymes), or it could be that these strains are not capable of degrading phytate.

## 6.5 Discussion

The work presented in this supplementary chapter shows a diverse community of *Burkholderiaceae* genera can be isolated from the roots of wheat. Long-read sequencing however showed that none of these genera are dominant within the endosphere or rhizosphere, with the exception of *Massilia* (Chapter Four, Section 4.5). This chapter has presented some interesting bioactivity and for *Comamonas* isolates, and antagonistic interactions between *Comamonas*, *Achromobacter*, or *Ralstonia* isolates and *Streptomyces* CRS3. No interesting activity was observed for *Massilia* however, the only one of these isolates shown to be prevalent within the endosphere and rhizosphere of wheat. For *Pseudomonas* there is a similar issue, whilst the *P. brassicacaerum* isolates showed activity against the wheat take-all fungus chapter four showed that it was the species *Pseudomonas poae* was present within the endosphere and rhizosphere. Long read sequencing demonstrated the redundancy of much of this culture-based work and exemplifies the core issues with culture-based work that was not informed by strain to genus level identifications for core microbiota within the root. Has this knowledge been acquired prior to the instigation of the work presented in this chapter, a more targeted approach could have been taken, with the use of media specifically targeting *Massilia* and *Duganella*, or a targeted screen to assay a larger number of

isolates in order to find the genera and species shown to be prevalent within the root. Regardless, this chapter has presented some interesting data relating to the ability of a cohort of root endosphere isolates to interact with *Streptomyces* spp. and to exhibit bioactivity against indicator strains and against the wheat take-all fungus.

## Chapter 7. Conclusions and future work

### 7.1 Conclusions

This project aimed to address four key questions, and to contribute significant insights into our understanding of key taxa associated with the root associated microbiome of UK bread wheat.

1. Can we detect any core bacterial, fungal, or archaeal taxa associated with UK bread wheat?

No selection for specific archaeal lineages was detected within the root, this was attributed to the fact that archaeal sequencing databases are incomplete and lack the established framework of their bacterial counterpart<sup>251</sup>. For the fungal community two taxa were found across multiple soil types, *Mortierellaceae* were associated with the rhizosphere and *Leotiaceae* were associated with the endosphere. As no data was acquired from the endosphere for plants cultivated in Levingtons compost, it is unclear if these fungi can be considered core root-associated taxa. The bacterial community was assessed across a range of genotypes in addition to soil types, five core associated taxa were found to be enriched with the root microbiome across all conditions tested, *Streptomycetaceae*, *Pseudomonadaceae*, *Burkholderiaceae*, *Rhizobiaceae*, and *Chitinophagaceae*. The identification of *Pseudomonadaceae*, *Burkholderiaceae*, *Rhizobiaceae*, and *Chitinophagaceae* corroborates the identification of these taxa within published studies as core microbes associated with wheat. Long-read sequencing of the 16S rRNA gene revealed specific genera and species within these families which are hypothesised to be associated with the wheat root. These included *Streptomyces turgidiscabies*, *Pseudomonas poae*, *Massilia* and *Duganella*, *Rhizobium* and *Agrobacterium*, and *Niastella* and *Chitinophaga* for each family respectively. This provides some confident associations between these families and UK wheat varieties, and some specific genera which are likely to interact with wheat. Future work can target these families and genera when investigating how these core microbes interact with the wheat root.

2. How do key factors such as genotype, soil type, and developmental senescence effect archaeal, bacterial, and fungal microbiome composition?

Despite the identification of core microbial taxa associated with the root microbiome, this work found that soil type had a significant effect on the microbial community,

concurrent with well-established literature demonstrating this principle<sup>52–56</sup>.

Genotype was not found to significantly affect the bacterial community composition, whilst this agrees with some published work<sup>128,129</sup>, it contradicts some others<sup>97,109</sup>.

These experiments contribute to ongoing bodies of work investigating factors which influence the microbiome and justifies future work comprehensively investigating these factors using longer read sequencing methodologies or using more standardised methodology.

Developmental senescence was found to have a profound effect on the microbial community; bacteria, fungi, and archaea all increased in abundance within the endosphere after plants has senesced, and the abundance of all of the core root associated families was significantly reduced. This indicated that there may be input from the living plant required to maintain these groups within the roots and warrants future work investigating microbiome dynamics during developmental senescence.

3. Which of these core microbes, if any, are able to utilise host derived carbon within root associated niches?

A DNA stable isotope probing experiment was able to show that nine different bacterial taxa within the rhizosphere were able to utilise host derived carbon. This included a number of the core root associated taxa such as *Oxalobacteraceae* and *Commamonadaceae* (within the *Burkholderiaceae*), and *Pseudomonadaceae*. Long read amplicon sequencing was able to show which bacterial genera or species within these families were most likely to be utilising host derived carbon.

Surprisingly *Streptomycetaceae*, the most abundant core associated taxon, was not utilising root exudates. There are a number of possible explanations for this, including that for DNA stable isotope probing experiments slower growing organisms can be more difficult to detect, particularly organisms such as *Streptomyces* which grow via hyphal tip extension and do not replicate their DNA frequently. Regardless, this experiment has provided intriguing insights into which microbiota interact most closely within the plant and raises interesting future questions surrounding the nature of those interactions.

4. What role do ammonia oxidising archaea play within the root?

As has been mentioned, no selection was detected for specific archaeal lineages within the root, nor were archaea found to be utilising root exudates. A culture-based experiment also showed that treating plants with a *Nitrososcosmicus* genus ammonia oxidising archaeon did not have a significant impact on the growth of wheat. DGGE analysis of archaeal 16S rRNA gene and *amoA* gene diversity

however indicated that there was a shift in the archaeal community across root compartments, and specifically within the community of ammonia oxidising archaea. This demonstrates that wheat roots can indeed influence the archaeal community within the roots, and specifically the community of AOA. The nature of this influence, and what it means for the host plant, remains a mystery.

Overall, this project has significantly contributed to understanding of the wheat root associated microbiome. Using metabarcoding the effect of genotype, soil type, and specific developmental stages has been investigated. This has contributed significantly to our understanding of the factors which can influence the root microbiome of wheat, and how this microbiome changes at the end of the plant's life. A number of core root associated microbes for wheat have been identified, and some of these have been identified to the genus/species level through long-read sequencing. Stable isotope probing has also shown a number of these taxa utilise host derived carbon, providing indication of the carbon source utilised by these organisms within the rhizosphere, and evidence for a close association between these taxa and the host. Cumulatively these results have provided a solid basis for future studies aiming to identify the molecular basis for interactions between key wheat root community members and the host plant, and the functional capacity of these organisms to provide services to the host. In the future this understanding may aid in the development of new strategies and microbial biotechnologies that can address many of the key issues facing agriculture in the 21<sup>st</sup> century.

## **7.2 Future work**

The work presented in this thesis has raised a number of questions that could be addressed in the future. Five core root associated fungal and bacterial taxa have been identified in association with roots, yet only a small number of these have been isolated and characterised. Future work could seek to culture these organisms from the roots of wheat, perhaps using root exudates or crushed root material as a carbon or nitrogen source for the isolation of new strains, or aiming to use co-cultivation techniques, directed evolution, or iCHIP enrichment to isolate difficult to cultivate endophytes.

Developmental senescence has also been shown to negatively impact the abundance of core root associated taxa. Future work could perform microscopy using labelled model necrotrophic fungi and bacteria to probe how these microbes colonise the root after senescence, and to comprehensively track the succession process that occurs after senescence. Further an experiment could fluorescently

label core root associated microbiota to image the fate of these microbes after senescence. RNA sequencing could also be performed to investigate how gene expression within core taxa such as *Streptomyces* changes within the root as the plant senesces, and to investigate why these microbes are no longer able to persist within the root after senescence.

Long read amplicon sequencing was able to identify a number of bacterial taxa within the endosphere to the genus/species level, however for most samples a low number of good quality reads were recovered, and this experiment failed to provide any information on the seed community. This experiment could be repeated with a greater sequencing depth to provide greater coverage for the microbial community. In addition to the PNA blockers used here, host depletion methods could also be used to further reduce host contamination and increase sequence depth for the microbial community. Metagenomics was also attempted on the endosphere of wheat within this project (data not included). Whilst this failed to provide sufficient coverage of the microbial community for the recovery of metagenome assembled genomes, or for any community level analysis, host cell depletion methods were able to reduce the extent of contamination by host derived sequences. In the future this method could be developed further, and deeper sequencing could be used to attempt to perform metagenomics from the endosphere.

For the archaeal community no major selection of taxa was detected across root compartments, despite DGGE indicating a shift in community diversity in root compartments compared to the bulk soil. Longer read archaeal 16S rRNA gene sequencing or full-length sequencing of the *amoA* gene could be used to better understand how archaeal diversity changes across root compartments, AOA diversity in particular could be investigated using the database produced by Alves *et al.*<sup>251</sup>. Enrichment culture was attempted within this project, however this failed to produce any consistent AOA enrichments, or to isolate any AOA strains (data not included). This experiment could be repeated, testing a broader variety of enrichment conditions, or using a crushed-exudate enrichment medium as was used by Simon *et al.*<sup>263</sup>. To follow up archaeal plant growth promotion investigations, this experiment could be repeated using an archaeal isolate acquired from the roots, or a different archaeal strain which is active under ambient conditions. Plants could also be cultivated at a higher temperature, and the effect of archaeal treatments on different plants or under different conditions (such as salt stress or nitrogen starvation) could also be tested.

To follow up the work in chapter four and validate the status of the five bacterial taxa identified as core root associated microbes, multiple varieties of wheat can be cultivated in a broader variety of soil types. In particular this can help to test how common the interaction between *Streptomyces* and wheat is, as the identification of *Streptomyces* is unique to this work and this taxon has not been proposed as a core root taxon by many published studies<sup>97,109,128</sup>. Longer read sequencing, for example using PacBio as mentioned, can also be used to identify strain-level changes in community composition across the root compartments and yield deeper insights into the identity of the core root associated community, and to understand whether changes in the composition of individual species within those core root families corroborates the status of those families as core root associated taxa.

Chapter five showed that only some core root associated microbes are capable of utilising host-derived carbon in the rhizosphere. To probe slower growing species more accurately RNA SIP should be used in the future to investigate whether streptomycetes or fungi within the rhizosphere and endosphere are active and able to utilise host derived carbon. Metagenomics or metatranscriptomics could also be applied to identify microbial functions selected for within the rhizosphere by root exudates. RNA SIP could be applied to the endosphere compartment, for which the work presented here failed to accurately assess root exudate utilisation. In addition, more replicates could also be used to account for greater variation within the endosphere microbiome. To probe other modes of host-derived carbon utilisation, such as direct tissue degradation, or acquisition of carbon from cell deposition within the rhizosphere, a longer labelling period could be used. Labelling plants for over a month would ensure that host tissues become sufficiently labelled and that microorganisms utilizing tissue derived carbon would become labelled in addition to those which utilise root exudates. One caveat from such an experiment is that organisms labelled via cross-feeding would be indistinguishable from those utilising tissue derived carbon. To overcome this, plants could be transplanted to fresh soil, which has not been exposed to <sup>13</sup>CO<sub>2</sub>, though additional controls would be required to assess the extent to which this process could stress the plant and disrupt the root microbiome. Further, this would not eradicate cross feeding-labelled microbiota within the endosphere compartment.

To follow up the identification of exudate utilising organisms targeted isolation approaches could be used to culture representatives of those genera or species from the endosphere. Then, using a hydroponic system to harvest root exudates from wheat at scale, culture-based experiments coupled with gas chromatography-

mass spectrometry (GC-MS) to identify the root exudate compounds which these organisms can use as a growth substrate. The effect of root exudates on the microorganism's behaviour and gene expression can also be monitored to further characterise the nature of the interaction between these core root exudate utilising microbes and the host.

In summary this project has produced a number of interesting biological questions, justifying a range of future investigations. Of most interest is the role of *Streptomyces* species within the root endosphere; this taxon is not always commonly identified in the literature yet was enriched within all conditions tested in this work. Key experiments to probe the interactions between this group and the host organism include RNA SIP, and microscopy experiments to visualise how *Streptomyces* spp. colonise the root. Further characterising how the other core microbiota interact with the host would be another key follow up study to this work, and could be achieved by combining isolation experiments, *in planta* assays for growth promotion and disease experiments, and investigations into exudate utilisation by these organisms. Lastly, further work can be done to validate the core-taxa status of core root associated families of bacteria by surveying the community in a broader variety of conditions and using longer read sequencing methods. Overall, this project has provided a solid base of evidence to justify these specific future investigations into the microbial ecology of the root associated microbiome of wheat.

Such investigations could yield knowledge that will inform new agricultural strategies such as crop breeding programs to select for genotypes that can maximise the beneficial microbiota harboured within the roots. Genetic investigations could identify host gene markers, or specific genome engineering interventions that could provide further tools to breeders aiming to generate strains with a strong beneficial root community. Similarly, genetic study of beneficial microbiota can understand the mechanisms used by beneficial microbiota for the colonisation of roots and use this information to engineer beneficial strains or synthetic communities that can consistently colonise roots under a wide range of conditions. Microbial biotechnologies can be developed by applying the identified core root associated taxa as, for example, biocontrol formulations. Isolates described in chapter 6 reinforce this possibility as they demonstrate the capacity to impede the wheat take-all fungus. This work has fed into a large range of innovations and scientific avenues that can enhance our understanding of plant-microbe interactions, and how

these interactions can be used to enhance agricultural productivity as the climate changes, and we shift to more ecologically responsible means of food production

## Bibliography

1. Zeder, M. A. The Origins of Agriculture in the Near East. *Current Anthropology* **52**, S221–S235 (2011).
2. Snir, A. *et al.* The Origin of Cultivation and Proto-Weeds, Long Before Neolithic Farming. *PLoS ONE* **10**, e0131422 (2015).
3. Tilman, D., Cassman, K. G., Matson, P. A., Naylor, R. & Polasky, S. Agricultural sustainability and intensive production practices. *Nature* **418**, 671–677 (2002).
4. Montgomery, D. R. Soil erosion and agricultural sustainability. *Proceedings of the National Academy of Sciences* **104**, 13268–13272 (2007).
5. Nishida, M. Decline in Fertility of Paddy Soils Induced by Paddy Rice and Upland Soybean Rotation, and Measures against the Decline. *JARQ* **50**, 87–94 (2016).
6. Elias, E. & Scoones, I. Perspectives on soil fertility change: a case study from southern Ethiopia. *LAND DEGRADATION* **12** (1999).
7. Vergé, X. P. C., De Kimpe, C. & Desjardins, R. L. Agricultural production, greenhouse gas emissions and mitigation potential. *Agricultural and Forest Meteorology* **142**, 255–269 (2007).
8. Tschardtke, T., Klein, A. M., Krüss, A., Steffan-Dewenter, I. & Thies, C. Landscape perspectives on agricultural intensification and biodiversity – ecosystem service management. *Ecology Letters* **8**, 857–874 (2005).
9. Wheeler, R. & Lobley, M. Managing extreme weather and climate change in UK agriculture: Impacts, attitudes and action among farmers and stakeholders. *Climate Risk Management* **32**, 100313 (2021).
10. Borrill, P., Harrington, S. A. & Uauy, C. Applying the latest advances in genomics and phenomics for trait discovery in polyploid wheat. *Plant J* **tpj.14150** (2018) doi:10.1111/tpj.14150.
11. Myers, S. S. *et al.* Climate Change and Global Food Systems: Potential Impacts on Food Security and Undernutrition. *Annu. Rev. Public Health* **38**, 259–277 (2017).
12. Gruber, N. & Galloway, J. N. An Earth-system perspective of the global nitrogen cycle. *Nature* **451**, 293–296 (2008).
13. Ashley, K., Cordell, D. & Mavinic, D. A brief history of phosphorus: From the philosopher's stone to nutrient recovery and reuse. *Chemosphere* **84**, 737–746 (2011).
14. Cordell, D., Drangert, J.-O. & White, S. The story of phosphorus: Global food security and food for thought. *Global Environmental Change* **19**, 292–305 (2009).
15. Potts, S. G. *et al.* Global pollinator declines: trends, impacts and drivers. *Trends in Ecology & Evolution* **25**, 345–353 (2010).
16. Goulson, D., Nicholls, E., Botias, C. & Rotheray, E. L. Bee declines driven by combined stress from parasites, pesticides, and lack of flowers. *Science* **347**, 1255957–1255957 (2015).
17. Gavrilescu, M. Fate of Pesticides in the Environment and its Bioremediation. *Eng. Life Sci.* **5**, 497–526 (2005).
18. Verweij, P. E., Snelders, E., Kema, G. H., Mellado, E. & Melchers, W. J. Azole resistance in *Aspergillus fumigatus*: a side-effect of environmental fungicide use? *The Lancet Infectious Diseases* **9**, 789–795 (2009).
19. Ma, Z. & Michailides, T. J. Advances in understanding molecular mechanisms of fungicide resistance and molecular detection of resistant genotypes in phytopathogenic fungi. *Crop Protection* **24**, 853–863 (2005).
20. Sangster, N. C. Anthelmintic resistance: past, present and future. *International Journal for Parasitology* **29**, 115–124 (1999).
21. Savary, S. *et al.* The global burden of pathogens and pests on major food crops. *Nat Ecol Evol* **3**, 430–439 (2019).
22. Ali, S. *et al.* Yellow Rust Epidemics Worldwide Were Caused by Pathogen Races from Divergent Genetic Lineages. *Front. Plant Sci.* **8**, 1057 (2017).

23. Milus, E. A., Kristensen, K. & Hovmøller, M. S. Evidence for Increased Aggressiveness in a Recent Widespread Strain of *Puccinia striiformis* f. sp. *tritici* Causing Stripe Rust of Wheat. *Phytopathology* **99**, 89–94 (2009).
24. United Nations, Department of Economic and Social Affairs PD. *World Population Prospects The 2017 Revision Key Findings and Advance Tables*. 1–46 (2017).
25. Fitzpatrick, C. R. *et al.* Assembly and ecological function of the root microbiome across angiosperm plant species. *Proc Natl Acad Sci USA* **115**, E1157–E1165 (2018).
26. Yandigeri, M. S. *et al.* Drought-tolerant endophytic actinobacteria promote growth of wheat (*Triticum aestivum*) under water stress conditions. *Plant Growth Regul* **68**, 411–420 (2012).
27. Farwell, A. J. *et al.* Tolerance of transgenic canola plants (*Brassica napus*) amended with plant growth-promoting bacteria to flooding stress at a metal-contaminated field site. *Environmental Pollution* **147**, 540–545 (2007).
28. Fan, P. *et al.* Alleviating salt stress in tomato seedlings using *Arthrobacter* and *Bacillus megaterium* isolated from the rhizosphere of wild plants grown on saline–alkaline lands. *International Journal of Phytoremediation* **18**, 1113–1121 (2016).
29. Palaniyandi, S. A., Damodharan, K., Yang, S. H. & Suh, J. W. *Streptomyces* sp. strain PGPA39 alleviates salt stress and promotes growth of ‘Micro Tom’ tomato plants. *J Appl Microbiol* **117**, 766–773 (2014).
30. Khalofah, A., Kilany, M. & Migdadi, H. Phytostimulatory Influence of *Comamonas testosteroni* and Silver Nanoparticles on *Linum usitatissimum* L. under Salinity Stress. *Plants* **10**, 790 (2021).
31. Jaiswal, S. K., Mohammed, M., Ibny, F. Y. I. & Dakora, F. D. Rhizobia as a Source of Plant Growth-Promoting Molecules: Potential Applications and Possible Operational Mechanisms. *Front. Sustain. Food Syst.* **4**, 619676 (2021).
32. Jog, R., Pandya, M., Nareshkumar, G. & Rajkumar, S. Mechanism of phosphate solubilization and antifungal activity of *Streptomyces* spp. isolated from wheat roots and rhizosphere and their application in improving plant growth. *Microbiology* **160**, 778–788 (2014).
33. Fahsi, N., Mahdi, I., Mesfioui, A., Biskri, L. & Allaoui, A. Phosphate solubilizing rhizobacteria isolated from jujube *Ziziphus lotus* plant stimulate wheat germination rate and seedlings growth. *PeerJ* **9**, e11583 (2021).
34. Püschel, D., Bitterlich, M., Rydlová, J. & Jansa, J. Drought accentuates the role of mycorrhiza in phosphorus uptake. *Soil Biology and Biochemistry* **157**, 108243 (2021).
35. Sah, S., Singh, N. & Singh, R. Iron acquisition in maize (*Zea mays* L.) using *Pseudomonas* siderophore. *3 Biotech* **7**, 121 (2017).
36. Balbino, H. M., Monteiro, T. S. A., Coutinho, R. R., Pacheco, P. V. M. & Freitas, L. G. de. Association of *Duddingtonia flagrans* with microorganisms for management of *Meloidogyne javanica* and acquisition of nutrients in soybean. *Biological Control* **159**, 104626 (2021).
37. Carrión, V. J. *et al.* Involvement of *Burkholderiaceae* and sulfurous volatiles in disease-suppressive soils. *ISME J* **12**, 2307–2321 (2018).
38. Gómez Expósito, R., de Bruijn, I., Postma, J. & Raaijmakers, J. M. Current Insights into the Role of Rhizosphere Bacteria in Disease Suppressive Soils. *Front. Microbiol.* **8**, 2529 (2017).
39. Pane, C., Spaccini, R., Piccolo, A., Scala, F. & Bonanomi, G. Compost amendments enhance peat suppressiveness to *Pythium ultimum*, *Rhizoctonia solani* and *Sclerotinia minor*. *Biological Control* **56**, 115–124 (2011).
40. Yang, M., Mavrodi, D. V., Thomashow, L. S. & Weller, D. M. Differential Response of Wheat Cultivars to *Pseudomonas brassicacearum* and Take-All Decline Soil. *Phytopathology* **108**, 1363–1372 (2018).
41. Worsley, S. F. *et al.* *Streptomyces* Endophytes Promote Host Health and Enhance Growth across Plant Species. *Appl Environ Microbiol* **86**, e01053-20, /aem/86/16/AEM.01053-20.atom (2020).
42. Shariffah-Muzaimah, S. A. *et al.* Characterization of *Streptomyces* spp. isolated from the rhizosphere of oil palm and evaluation of their ability to suppress basal stem rot disease in oil palm seedlings when applied as powder formulations in a glasshouse trial. *World J Microbiol Biotechnol* **34**, 15 (2018).

43. Yu, W. Q. *et al.* *Paenibacillus terrae* NK3-4: A potential biocontrol agent that produces  $\beta$ -1,3-glucanase. *Biological Control* **129**, 92–101 (2019).
44. Preston, G. M. Plant perceptions of plant growth-promoting *Pseudomonas*. *12* (2004).
45. Lahdenperä, M. L., Simon, E. & Uoti, J. Mycostop-A Novel Biofungicide Based on *Streptomyces* Bacteria. *Developments in Agricultural and Managed Forest Ecology* **23**, 258–263 (1991).
46. Himmelstein, J. C., Maul, J. E. & Everts, K. L. Impact of Five Cover Crop Green Manures and Actinovate on *Fusarium* Wilt of Watermelon. *Plant Disease* **98**, 965–972 (2014).
47. Sabaratnam, S. & Traquair, J. A. Mechanism of antagonism by *Streptomyces griseocarneus* (strain Di944) against fungal pathogens of greenhouse-grown tomato transplants. *Canadian Journal of Plant Pathology* **37**, 197–211 (2015).
48. Tokala, R. K. *et al.* Novel Plant-Microbe Rhizosphere Interaction Involving *Streptomyces lydicus* WYEC108 and the Pea Plant (*Pisum sativum*). *Appl Environ Microbiol* **68**, 2161–2171 (2002).
49. Chaparro, J. M., Badri, D. V. & Vivanco, J. M. Rhizosphere microbiome assemblage is affected by plant development. *ISME J* **8**, 790–803 (2014).
50. Edwards, J. *et al.* Structure, variation, and assembly of the root-associated microbiomes of rice. *Proc Natl Acad Sci USA* **112**, E911–E920 (2015).
51. Bulgarelli, D. *et al.* Structure and Function of the Bacterial Root Microbiota in Wild and Domesticated Barley. *Cell Host & Microbe* **17**, 392–403 (2015).
52. Bulgarelli, D. *et al.* Revealing structure and assembly cues for *Arabidopsis* root-inhabiting bacterial microbiota. *Nature* **488**, 91–95 (2012).
53. Rousk, J. *et al.* Soil bacterial and fungal communities across a pH gradient in an arable soil. *ISME J* **4**, 1340–1351 (2010).
54. Pershina, E. V. *et al.* Investigation of the core microbiome in main soil types from the East European plain. *Science of The Total Environment* **631–632**, 1421–1430 (2018).
55. Berg, G. & Smalla, K. Plant species and soil type cooperatively shape the structure and function of microbial communities in the rhizosphere: Plant species, soil type and rhizosphere communities. *FEMS Microbiology Ecology* **68**, 1–13 (2009).
56. Marschner, P. Soil and plant specific effects on bacterial community composition in the rhizosphere. *Soil Biology* **9** (2001).
57. Seang-On, L. Alleviation of cadmium stress in thai rice cultivar (psl2) by inoculation of indigenous cadmium resistant microbial consortia. *Appl. Ecol. Env. Res.* **17**, (2019).
58. Carvalhais, L. C., Muzzi, F., Tan, C.-H., Hsien-Choo, J. & Schenk, P. M. Plant growth in *Arabidopsis* is assisted by compost soil-derived microbial communities. *Front. Plant Sci.* **4**, (2013).
59. Hugoni, M., Luis, P., Guyonnet, J. & Haichar, F. el Z. Plant host habitat and root exudates shape fungal diversity. *Mycorrhiza* **28**, 451–463 (2018).
60. Garbeva, P., van Veen, J. A. & van Elsas, J. D. Microbial diversity in soil: selection of microbial populations by plant and soil type and implications for disease suppressiveness. *Annu. Rev. Phytopathol.* **42**, 243–270 (2004).
61. Miller, C. J., Yesiller, N., Yaldo, K. & Merayyan, S. Impact of Soil Type and Compaction Conditions on Soil Water Characteristic. *J. Geotech. Geoenviron. Eng.* **128**, 733–742 (2002).
62. Grady, K. L., Sorensen, J. W., Stopnisek, N., Guittar, J. & Shade, A. Assembly and seasonality of core phyllosphere microbiota on perennial biofuel crops. *Nat Commun* **10**, 4135 (2019).
63. Mavrodi, D. V. *et al.* Long-Term Irrigation Affects the Dynamics and Activity of the Wheat Rhizosphere Microbiome. *Front. Plant Sci.* **9**, 345 (2018).
64. Paul Chowdhury, S. *et al.* Effect of long-term organic and mineral fertilization strategies on rhizosphere microbiota assemblage and performance of lettuce. *Environ Microbiol* **21**, 2426–2439 (2019).
65. Ai, C. *et al.* Reduced dependence of rhizosphere microbiome on plant-derived carbon in 32-year long-term inorganic and organic fertilized soils. *Soil Biology and Biochemistry* **80**, 70–78 (2015).

66. Wang, Q. *et al.* Long-term N fertilization altered <sup>13</sup>C-labeled fungal community composition but not diversity in wheat rhizosphere of Chinese black soil. *Soil Biology and Biochemistry* **135**, 117–126 (2019).
67. Babin, D. *et al.* Impact of long-term agricultural management practices on soil prokaryotic communities. *Soil Biology and Biochemistry* **129**, 17–28 (2019).
68. Turner, T. R. *et al.* Comparative metatranscriptomics reveals kingdom level changes in the rhizosphere microbiome of plants. *ISME J* **7**, 2248–2258 (2013).
69. Bressan, M. *et al.* Exogenous glucosinolate produced by *Arabidopsis thaliana* has an impact on microbes in the rhizosphere and plant roots. *ISME J* **3**, 1243–1257 (2009).
70. Sawers, R. J. H., Ramírez-Flores, M. R., Olalde-Portugal, V. & Paszkowski, U. The impact of domestication and crop improvement on arbuscular mycorrhizal symbiosis in cereals: insights from genetics and genomics. *New Phytol* **220**, 1135–1140 (2018).
71. Liu, A., Ku, Y.-S., Contador, C. A. & Lam, H.-M. The Impacts of Domestication and Agricultural Practices on Legume Nutrient Acquisition Through Symbiosis With Rhizobia and Arbuscular Mycorrhizal Fungi. *Front. Genet.* **11**, 583954 (2020).
72. Iannucci, A., Fragasso, M., Beleggia, R., Nigro, F. & Papa, R. Evolution of the Crop Rhizosphere: Impact of Domestication on Root Exudates in Tetraploid Wheat (*Triticum turgidum* L.). *Front. Plant Sci.* **8**, 2124 (2017).
73. Chaparro, J. M. *et al.* Root Exudation of Phytochemicals in *Arabidopsis* Follows Specific Patterns That Are Developmentally Programmed and Correlate with Soil Microbial Functions. *PLOS ONE* **8**, 10 (2013).
74. Micallef, S. A., Shiaris, M. P. & Colón-Carmona, A. Influence of *Arabidopsis thaliana* accessions on rhizobacterial communities and natural variation in root exudates. *Journal of Experimental Botany* **60**, 1729–1742 (2009).
75. Ou, T. *et al.* A Microbiome Study Reveals Seasonal Variation in Endophytic Bacteria Among different Mulberry Cultivars. *Computational and Structural Biotechnology Journal* **17**, 1091–1100 (2019).
76. Hu, W., Strom, N. B., Haarith, D., Chen, S. & Bushley, K. E. Seasonal Variation and Crop Sequences Shape the Structure of Bacterial Communities in Cysts of Soybean Cyst Nematode. *Front. Microbiol.* **10**, 2671 (2019).
77. Agler, M. T. *et al.* Microbial Hub Taxa Link Host and Abiotic Factors to Plant Microbiome Variation. *PLoS Biol* **14**, e1002352 (2016).
78. Marsh, P. & Wellington, E. M. H. Phage-host interactions in soil. *FEMS Microbiology Ecology* **15**, 99–107 (1994).
79. Sutela, S., Poimala, A. & Vainio, E. J. Viruses of fungi and oomycetes in the soil environment. *FEMS Microbiology Ecology* **95**, f1119 (2019).
80. Ghanem, N., Stanley, C. E., Harms, H., Chatzinotas, A. & Wick, L. Y. Mycelial Effects on Phage Retention during Transport in a Microfluidic Platform. *Environ. Sci. Technol.* **53**, 11755–11763 (2019).
81. Yaegashi, H. *et al.* Appearance of mycovirus-like double-stranded RNAs in the white root rot fungus, *Rosellinia necatrix*, in an apple orchard. *FEMS Microbiol Ecol* **83**, 49–62 (2013).
82. Torres-Trenas, A., Prieto, P., Cañizares, M. C., García-Pedrajas, M. D. & Pérez-Artés, E. Mycovirus *Fusarium oxysporum* f. sp. *dianthi* Virus 1 Decreases the Colonizing Efficiency of Its Fungal Host. *Front. Cell. Infect. Microbiol.* **9**, 51 (2019).
83. Neupane, A. *et al.* Metatranscriptomic Analysis and In Silico Approach Identified Mycoviruses in the Arbuscular Mycorrhizal Fungus *Rhizophagus* spp. *Viruses* **10**, 707 (2018).
84. Garbeva, P., van Elsas, J. D. & van Veen, J. A. Rhizosphere microbial community and its response to plant species and soil history. *Plant Soil* **302**, 19–32 (2008).
85. Sasse, J., Martinoia, E. & Northen, T. Feed Your Friends: Do Plant Exudates Shape the Root Microbiome? *Trends in Plant Science* **23**, 25–41 (2018).

86. Satoh, K. *et al.* Effect of root exudates on growth of newly isolated nitrifying bacteria from barley rhizosphere. *Soil Science and Plant Nutrition* **49**, 757–762 (2003).
87. Tsavkelova, E. A., Cherdynitseva, T. A., Lobakova, E. S., Kolomeitseva, G. L. & Netrusov, A. I. Microbiota of the Orchid Rhizosphere. **70**, 6 (2001).
88. Kurakov, A. V. & Kostina, N. V. Spatial Peculiarities in the Colonization of the Plant Rhizosphere by Microscopic Fungi. *Aspergillus fumigatus* **70**, 10 (2001).
89. Sadeghi, A. *et al.* Plant growth promoting activity of an auxin and siderophore producing isolate of *Streptomyces* under saline soil conditions. *World J Microbiol Biotechnol* **28**, 1503–1509 (2012).
90. Zia, M. A. *et al.* Glucanolytic rhizobacteria associated with wheat- maize cropping system suppress the Fusarium wilt of tomato (*Lycopersicon esculentum* L.). *Scientia Horticulturae* **287**, 110275 (2021).
91. Vandenkoornhuysen, P., Quaiser, A., Duhamel, M., Le Van, A. & Dufresne, A. The importance of the microbiome of the plant holobiont. *New Phytol* **206**, 1196–1206 (2015).
92. Bragina, A., Berg, C. & Berg, G. The core microbiome binds the Alpine bog vegetation to a transkingdom metacommunity. *Mol Ecol* **24**, 4795–4807 (2015).
93. Shade, A. & Handelsman, J. Beyond the Venn diagram: the hunt for a core microbiome: The hunt for a core microbiome. *Environmental Microbiology* **14**, 4–12 (2012).
94. Rascovan, N. *et al.* Integrated analysis of root microbiomes of soybean and wheat from agricultural fields. *Sci Rep* **6**, 28084 (2016).
95. Houlden, A., Timms-Wilson, T. M., Day, M. J. & Bailey, M. J. Influence of plant developmental stage on microbial community structure and activity in the rhizosphere of three field crops: Plant and growth stage effects on microbial populations. *FEMS Microbiology Ecology* **65**, 193–201 (2008).
96. Chen, S. *et al.* Root-associated microbiomes of wheat under the combined effect of plant development and nitrogen fertilization. *Microbiome* **7**, 136 (2019).
97. Tkacz, A. *et al.* Agricultural Selection of Wheat Has Been Shaped by Plant-Microbe Interactions. *Front. Microbiol.* **11**, 132 (2020).
98. Granzow, S. *et al.* The Effects of Cropping Regimes on Fungal and Bacterial Communities of Wheat and Faba Bean in a Greenhouse Pot Experiment Differ between Plant Species and Compartment. *Front. Microbiol.* **8**, 902 (2017).
99. Gdanetz, K. & Trail, F. The Wheat Microbiome Under Four Management Strategies, and Potential for Endophytes in Disease Protection. *Phytobiomes* **1**, 158–168 (2017).
100. Lundberg, D. S. *et al.* Defining the core *Arabidopsis thaliana* root microbiome. *Nature* **488**, 86–90 (2012).
101. Tian, B. *et al.* Beneficial traits of bacterial endophytes belonging to the core communities of the tomato root microbiome. *Agriculture, Ecosystems & Environment* **247**, 149–156 (2017).
102. Yeoh, Y. K. *et al.* The core root microbiome of sugarcane cultivated under varying nitrogen fertilizer application: N fertilizer and sugarcane root microbiota. *Environ Microbiol* **18**, 1338–1351 (2016).
103. Yeoh, Y. K. *et al.* Evolutionary conservation of a core root microbiome across plant phyla along a tropical soil chronosequence. *Nat Commun* **8**, 215 (2017).
104. Risely, A. Applying the core microbiome to understand host–microbe systems. *J Anim Ecol* **89**, 1549–1558 (2020).
105. Brown, S. P. *et al.* Scraping the bottom of the barrel: are rare high throughput sequences artifacts? *Fungal Ecology* **13**, 221–225 (2015).
106. Skopina, M. Yu., Vasileva, A. A., Pershina, E. V. & Pinevich, A. V. Diversity at low abundance: The phenomenon of the rare bacterial biosphere. *Microbiology* **85**, 272–282 (2016).
107. Sogin, M. L. *et al.* Microbial diversity in the deep sea and the underexplored “rare biosphere”. **6** (2006).

108. Kuźniar, A. *et al.* Culture-independent analysis of an endophytic core microbiome in two species of wheat: *Triticum aestivum* L. (cv. 'Hondia') and the first report of microbiota in *Triticum spelta* L. (cv. 'Rokosz'). *Systematic and Applied Microbiology* **43**, 126025 (2020).
109. Schlatter, D. C., Yin, C., Hulbert, S. & Paulitz, T. C. Core Rhizosphere Microbiomes of Dryland Wheat Are Influenced by Location and Land Use History. *Appl Environ Microbiol* **86**, e02135-19, /aem/86/5/AEM.02135-19.atom (2019).
110. Germida, J. J. & Siciliano, S. Taxonomic diversity of bacteria associated with the roots of modern, recent and ancient wheat cultivars. *Biology and Fertility of Soils* **33**, 410–415 (2001).
111. Mahoney, A. K., Yin, C. & Hulbert, S. H. Community Structure, Species Variation, and Potential Functions of Rhizosphere-Associated Bacteria of Different Winter Wheat (*Triticum aestivum*) Cultivars. *Front. Plant Sci.* **8**, (2017).
112. Donn, S., Kirkegaard, J. A., Perera, G., Richardson, A. E. & Watt, M. Evolution of bacterial communities in the wheat crop rhizosphere: Rhizosphere bacteria in field-grown intensive wheat crops. *Environ Microbiol* **17**, 610–621 (2015).
113. Fan, K. *et al.* Rhizosphere-associated bacterial network structure and spatial distribution differ significantly from bulk soil in wheat crop fields. *Soil Biology and Biochemistry* **113**, 275–284 (2017).
114. Iannucci, A., Canfora, L., Nigro, F., De Vita, P. & Beleggia, R. Relationships between root morphology, root exudate compounds and rhizosphere microbial community in durum wheat. *Applied Soil Ecology* **158**, 103781 (2021).
115. Hayden, H. L., Savin, K. W., Wadeson, J., Gupta, V. V. S. R. & Mele, P. M. Comparative Metatranscriptomics of Wheat Rhizosphere Microbiomes in Disease Suppressive and Non-suppressive Soils for *Rhizoctonia solani* AG8. *Front. Microbiol.* **9**, 859 (2018).
116. Marschner, P., Solaiman, Z. & Rengel, Z. Growth, phosphorus uptake, and rhizosphere microbial-community composition of a phosphorus-efficient wheat cultivar in soils differing in pH. *J. Plant Nutr. Soil Sci.* **168**, 343–351 (2005).
117. Durán, P. *et al.* Microbial Community Composition in Take-All Suppressing Soils. *Front. Microbiol.* **9**, 2198 (2018).
118. Wang, S. *et al.* Wheat Rhizosphere Metagenome Reveals Newfound Potential Soil Zn-Mobilizing Bacteria Contributing to Cultivars' Variation in Grain Zn Concentration. *Front. Microbiol.* **12**, 689855 (2021).
119. Zhou, Y. *et al.* The preceding root system drives the composition and function of the rhizosphere microbiome. *Genome Biol* **21**, 89 (2020).
120. Ofek, M., Voronov-Goldman, M., Hadar, Y. & Minz, D. Host signature effect on plant root-associated microbiomes revealed through analyses of resident vs. active communities: Host effect on plant root-associated microbiomes. *Environ Microbiol* **16**, 2157–2167 (2014).
121. Yin, C. *et al.* Bacterial Communities on Wheat Grown Under Long-Term Conventional Tillage and No-Till in the Pacific Northwest of the United States. *Phytobiomes Journal* **1**, 83–90 (2017).
122. Sanguin, H., Sarniguet, A., Gazengel, K., Moëgne-Loccoz, Y. & Grundmann, G. L. Rhizosphere bacterial communities associated with disease suppressiveness stages of take-all decline in wheat monoculture. *New Phytologist* **184**, 694–707 (2009).
123. Yin, C. *et al.* Role of Bacterial Communities in the Natural Suppression of *Rhizoctonia solani* Bare Patch Disease of Wheat (*Triticum aestivum* L.). *Appl. Environ. Microbiol.* **79**, 7428–7438 (2013).
124. Hamonts, K. *et al.* Effect of nitrogen and waterlogging on denitrifier gene abundance, community structure and activity in the rhizosphere of wheat. *FEMS Microbiol Ecol* **83**, 568–584 (2013).
125. Sachdev, D., Nema, P., Dhakephalkar, P., Zinjarde, S. & Chopade, B. Assessment of 16S rRNA gene-based phylogenetic diversity and promising plant growth-promoting traits of *Acinetobacter* community from the rhizosphere of wheat. *Microbiological Research* **165**, 627–638 (2010).
126. Tkacz, A., Cheema, J., Chandra, G., Grant, A. & Poole, P. S. Stability and succession of the rhizosphere microbiota depends upon plant type and soil composition. *ISME J* **9**, 2349–2359 (2015).
127. Banerjee, S. *et al.* Agricultural intensification reduces microbial network complexity and the abundance of keystone taxa in roots. *ISME J* **13**, 1722–1736 (2019).

128. Simonin, M. *et al.* Influence of plant genotype and soil on the wheat rhizosphere microbiome: evidences for a core microbiome across eight African and European soils. *FEMS Microbiology Ecology* **96**, fiae067 (2020).
129. Özkurt, E. *et al.* Seed-Derived Microbial Colonization of Wild Emmer and Domesticated Bread Wheat (*Triticum dicoccoides* and *T. aestivum*) Seedlings Shows Pronounced Differences in Overall Diversity and Composition. *mBio* **11**, (2020).
130. Revillini, D., Wilson, G. W. T., Miller, R. M., Lancione, R. & Johnson, N. C. Plant Diversity and Fertilizer Management Shape the Belowground Microbiome of Native Grass Bioenergy Feedstocks. *Front. Plant Sci.* **10**, 1018 (2019).
131. Di Bella, J. M., Bao, Y., Gloor, G. B., Burton, J. P. & Reid, G. High throughput sequencing methods and analysis for microbiome research. *Journal of Microbiological Methods* **95**, 401–414 (2013).
132. Yarza, P. *et al.* Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. *Nat Rev Microbiol* **12**, 635–645 (2014).
133. Liu, Y.-X. *et al.* A practical guide to amplicon and metagenomic analysis of microbiome data. *Protein Cell* **12**, 315–330 (2021).
134. Pollock, J., Glendinning, L., Wisedchanwet, T. & Watson, M. The Madness of Microbiome: Attempting To Find Consensus “Best Practice” for 16S Microbiome Studies. *Appl Environ Microbiol* **84**, (2018).
135. Silverman, J. D. *et al.* Measuring and Mitigating PCR Bias in Microbiome Data. <http://biorxiv.org/lookup/doi/10.1101/604025> (2019) doi:10.1101/604025.
136. Kavamura, V. N. *et al.* Land Management and Microbial Seed Load Effect on Rhizosphere and Endosphere Bacterial Community Assembly in Wheat. *Frontiers in Microbiology* **10**, 11 (2019).
137. Worsley, S. F. *et al.* Investigating the Role of Root Exudates in Recruiting *Streptomyces* Bacteria to the *Arabidopsis thaliana* Microbiome. *Front. Mol. Biosci.* **8**, 686110 (2021).
138. Haichar, F. el Z., Heulin, T., Guyonnet, J. P. & Achouak, W. Stable isotope probing of carbon flow in the plant holobiont. *Current Opinion in Biotechnology* **41**, 9–13 (2016).
139. Huang, A. C. *et al.* A specialized metabolic network selectively modulates *Arabidopsis* root microbiota. *Science* **364**, eaau6389 (2019).
140. Zhalina, K. *et al.* Dynamic root exudate chemistry and microbial substrate preferences drive patterns in rhizosphere microbial community assembly. *Nat Microbiol* **3**, 470–480 (2018).
141. Moe, L. A. Amino acids in the rhizosphere: From plants to microbes. *American Journal of Botany* **100**, 1692–1705 (2013).
142. Badri, D. V., Chaparro, J. M., Zhang, R., Shen, Q. & Vivanco, J. M. Application of Natural Blends of Phytochemicals Derived from the Root Exudates of *Arabidopsis* to the Soil Reveal That Phenolic-related Compounds Predominantly Modulate the Soil Microbiome. *J. Biol. Chem.* **288**, 4502–4512 (2013).
143. Proctor, C. & He, Y. Quantifying root extracts and exudates of sedge and shrub in relation to root morphology. *Soil Biology and Biochemistry* **114**, 168–180 (2017).
144. Wang, Y.-L. *et al.* Contrasting responses of root morphology and root-exuded organic acids to low phosphorus availability in three important food crops with divergent root traits. *AoB PLANTS* **7**, plv097 (2015).
145. Li, X., Dong, J., Chu, W., Chen, Y. & Duan, Z. The relationship between root exudation properties and root morphological traits of cucumber grown under different nitrogen supplies and atmospheric CO<sub>2</sub> concentrations. *Plant Soil* **425**, 415–432 (2018).
146. Hussain, H. A. *et al.* Effects of Arbuscular Mycorrhizal Fungi on Maize Growth, Root Colonization, and Root Exudates Varied with Inoculum and Application Method. *J Soil Sci Plant Nutr* **21**, 1577–1590 (2021).
147. Wang, P., Bi, S., Wang, S. & Ding, Q. Variation of Wheat Root Exudates under Aluminum Stress. *J. Agric. Food Chem.* **54**, 10040–10046 (2006).

148. Fan, X., Wen, X., Huang, F., Cai, Y. & Cai, K. Effects of silicon on morphology, ultrastructure and exudates of rice root under heavy metal stress. *Acta Physiol Plant* **38**, 197 (2016).
149. Xu, Q. *et al.* Potassium Improves Drought Stress Tolerance in Plants by Affecting Root Morphology, Root Exudates, and Microbial Diversity. *Metabolites* **11**, 131 (2021).
150. Tantriani *et al.* Metabolomic analysis of night-released soybean root exudates under high- and low-K conditions. *Plant Soil* **456**, 259–276 (2020).
151. Oburger, E. *et al.* Root exudation of phytosiderophores from soil-grown wheat. *New Phytol* **203**, 1161–1174 (2014).
152. Pei, J. *et al.* Different responses of root exudates to biochar application under elevated CO<sub>2</sub>. *Agriculture, Ecosystems & Environment* **301**, 107061 (2020).
153. Badri, D. V., Loyola-Vargas, V. M., Broeckling, C. D. & Vivanco, J. M. Root Secretion of Phytochemicals in Arabidopsis Is Predominantly Not Influenced by Diurnal Rhythms. *Molecular Plant* **3**, 491–498 (2010).
154. Archetti, M. *et al.* Let the Right One In: A Microeconomic Approach to Partner Choice in Mutualisms. *The American Naturalist* **177**, 75–85 (2011).
155. de Weert, S. *et al.* Flagella-Driven Chemotaxis Towards Exudate Components Is an Important Trait for Tomato Root Colonization by *Pseudomonas fluorescens*. 8.
156. Feng, H. *et al.* Chemotaxis of Beneficial Rhizobacteria to Root Exudates: The First Step towards Root–Microbe Rhizosphere Interactions. *IJMS* **22**, 6655 (2021).
157. Bowra, B. J. & Dilworth, M. J. Motility and Chemotaxis towards Sugars in *Rhizobium leguminosarum*. **126**, 231–235 (1981).
158. T. A., P. D. *et al.* Bacterial rhizosphere community profile at different growth stages of Umorok (*Capsicum chinense*) and its response to the root exudates. *Int Microbiol* **23**, 241–251 (2020).
159. Gupta Sood, S. Chemotactic response of plant-growth-promoting bacteria towards roots of vesicular-arbuscular mycorrhizal tomato plants. *FEMS Microbiology Ecology* **45**, 219–227 (2003).
160. Guo, Q. *et al.* The biocontrol agent *Streptomyces pactum* increases *Pseudomonas koreensis* populations in the rhizosphere by enhancing chemotaxis and biofilm formation. *Soil Biology and Biochemistry* **144**, 107755 (2020).
161. Worsley, S. F. *et al.* Investigating the role of root exudates in recruiting *Streptomyces* bacteria to the *Arabidopsis thaliana* root microbiome. <http://biorxiv.org/lookup/doi/10.1101/2020.09.09.290742> (2020) doi:10.1101/2020.09.09.290742.
162. Nagata, M. *et al.* Enhanced hyphal growth of arbuscular mycorrhizae by root exudates derived from high R/FR treated *Lotus japonicus*. *Plant Signaling & Behavior* **11**, e1187356 (2016).
163. Mavrodi, O. V. *et al.* Root Exudates Alter the Expression of Diverse Metabolic, Transport, Regulatory, and Stress Response Genes in Rhizosphere *Pseudomonas*. *Front. Microbiol.* **12**, 651282 (2021).
164. Fan, B. *et al.* Transcriptomic profiling of *Bacillus amyloliquefaciens* FZB42 in response to maize root exudates. *BMC Microbiol* **12**, 116 (2012).
165. Bücking, H. *et al.* Root exudates stimulate the uptake and metabolism of organic carbon in germinating spores of *Glomus intraradices*. *New Phytologist* **180**, 684–695 (2008).
166. Wang, H.-W. *et al.* Root endophyte-enhanced peanut-rhizobia interaction is associated with regulation of root exudates. *Microbiological Research* **250**, 126765 (2021).
167. Bell, C. A., Lilley, C. J., McCarthy, J., Atkinson, H. J. & Urwin, P. E. Plant-parasitic nematodes respond to root exudate signals with host-specific gene expression patterns. *PLoS Pathog* **15**, e1007503 (2019).
168. Karlsson, A. E., Johansson, T. & Bengtson, P. Archaeal abundance in relation to root and fungal exudation rates. *FEMS Microbiol Ecol* **80**, 305–311 (2012).
169. Mayer, Z., Juhász, Á. & Posta, K. Mycorrhizal Root Exudates Induce Changes in the Growth and Fumonisin Gene (*Fum1*) Expression of *Fusarium proliferatum*. *Agronomy* **9**, 291 (2019).

170. Kobayashi, A., Kim, M. J. & Kawazu, K. Uptake and Exudation of Phenolic Compounds by Wheat and Antimicrobial Components of the Root Exudate. *Zeitschrift für Naturforschung C* **51**, 527–533 (1996).
171. Lanoue, A. *et al.* De novo biosynthesis of defense root exudates in response to *Fusarium* attack in barley. *New Phytologist* **185**, 577–588 (2010).
172. Yuan, J. *et al.* Effect of phenolic acids from banana root exudates on root colonization and pathogen suppressive properties of *Bacillus amyloliquefaciens* NJN-6. *Biological Control* **125**, 131–137 (2018).
173. Nóbrega, F. M., Santos, I. S., Cunha, M. Da., Carvalho, A. O. & Gomes, V. M. Antimicrobial proteins from cowpea root exudates: inhibitory activity against *Fusarium oxysporum* and purification of a chitinase-like protein. *Plant Soil* **272**, 223–232 (2005).
174. Shaposhnikov, A. I., Shakhnazarova, V. Yu., Vishnevskaya, N. A., Borodina, E. V. & Strunnikova, O. K. Aromatic Carboxylic Acids in Barley-Root Exudates and Their Influence on the Growth of *Fusarium culmorum* and *Pseudomonas fluorescens*. *Appl Biochem Microbiol* **56**, 344–351 (2020).
175. Cupples, A. M. Contaminant-Degrading Microorganisms Identified Using Stable Isotope Probing. *Chem. Eng. Technol.* **39**, 1593–1603 (2016).
176. Imachi, H. *et al.* Isolation of an archaeon at the prokaryote–eukaryote interface. *Nature* **577**, 519–525 (2020).
177. Berry, D. *et al.* Host-compound foraging by intestinal microbiota revealed by single-cell stable isotope probing. *Proceedings of the National Academy of Sciences* **110**, 4720–4725 (2013).
178. Campana, S., Busch, K., Hentschel, U., Muyzer, G. & de Goeij, J. M. DNA-stable isotope probing (DNA-SIP) identifies marine sponge-associated bacteria actively utilizing dissolved organic matter (DOM). *Environ Microbiol* 1462-2920.15642 (2021) doi:10.1111/1462-2920.15642.
179. Godwin, S. *et al.* Investigation of the microbial metabolism of carbon dioxide and hydrogen in the kangaroo foregut by stable isotope probing. *ISME J* **8**, 1855–1865 (2014).
180. Cardini, U. *et al.* Chemosymbiotic bivalves contribute to the nitrogen budget of seagrass ecosystems. *ISME J* **13**, 3131–3134 (2019).
181. Radajewski, S. *et al.* Identification of active methylotroph populations in an acidic forest soil by stable isotope probing. *Microbiology* **148**, 2331-2343 (2002).
182. Xu, J., Jia, Z., Lin, X. & Feng, Y. DNA-based stable isotope probing identifies formate-metabolizing methanogenic archaea in paddy soil. *Microbiological Research* **202**, 36–42 (2017).
183. Dumont, M. G. & Murrell, J. C. Stable isotope probing — linking microbial identity to function. *Nat Rev Microbiol* **3**, 499–504 (2005).
184. Carrión, O. *et al.* Diversity of isoprene-degrading bacteria in phyllosphere and soil communities from a high isoprene-emitting environment: a Malaysian oil palm plantation. *Microbiome* **8**, 81 (2020).
185. Dumont, M. G., Pommerenke, B. & Casper, P. Using stable isotope probing to obtain a targeted metatranscriptome of aerobic methanotrophs in lake sediment: SIP-metatranscriptomics of methanotrophs. *Environmental Microbiology Reports* n/a-n/a (2013) doi:10.1111/1758-2229.12078.
186. Jehmlich, N., Vogt, C., Lünsmann, V., Richnow, H. H. & von Bergen, M. Protein-SIP in environmental studies. *Current Opinion in Biotechnology* **41**, 26–33 (2016).
187. Bernard, L. *et al.* Dynamics and identification of soil microbial populations actively assimilating carbon from <sup>13</sup>C-labelled wheat residue as estimated by DNA- and RNA-SIP techniques. *Environ Microbiol* **9**, 752–764 (2007).
188. Kaplan, H., Ratering, S., Felix-Henningsen, P. & Schnell, S. Stability of in situ immobilization of trace metals with different amendments revealed by microbial <sup>13</sup>C-labelled wheat root decomposition and efflux-mediated metal resistance of soil bacteria. *Science of The Total Environment* **659**, 1082–1089 (2019).
189. Neufeld, J. D. *et al.* DNA stable-isotope probing. *Nat Protoc* **2**, 860–866 (2007).

190. Haichar, F. el Z., Roncato, M.-A. & Achouak, W. Stable isotope probing of bacterial community structure and gene expression in the rhizosphere of *Arabidopsis thaliana*. *FEMS Microbiol Ecol* **81**, 291–302 (2012).
191. Lu, Y. In Situ Stable Isotope Probing of Methanogenic Archaea in the Rice Rhizosphere. *Science* **309**, 1088–1090 (2005).
192. Mayumi, D., Yoshimoto, T., Uchiyama, H., Nomura, N. & Nakajima-Kambe, T. Seasonal Change in Methanotrophic Diversity and Populations in a Rice Field Soil Assessed by DNA-Stable Isotope Probing and Quantitative Real-Time PCR. *Microb. Environ.* **25**, 156–163 (2010).
193. Hernández, M., Dumont, M. G., Yuan, Q. & Conrad, R. Different Bacterial Populations Associated with the Roots and Rhizosphere of Rice Incorporate Plant-Derived Carbon. *Appl Environ Microbiol* **81**, 2244–2253 (2015).
194. Haichar, F. el Z. *et al.* Plant host habitat and root exudates shape soil bacterial community structure. *ISME J* **2**, 1221–1230 (2008).
195. Wang, F., Shi, N., Jiang, R., Zhang, F. & Feng, G. In situ stable isotope probing of phosphate-solubilizing bacteria in the rhizosphere. *EXBOTJ* **67**, 1689–1701 (2016).
196. Uksa, M. *et al.* Bacteria utilizing plant-derived carbon in the rhizosphere of *Triticum aestivum* change in different depths of an arable soil: Spatial distribution of rhizosphere bacteria. *Environmental Microbiology Reports* **9**, 729–741 (2017).
197. Macey, M. C., Pratscher, J., Crombie, A. T. & Murrell, J. C. Impact of plants on the diversity and activity of methylophiles in soil. *Microbiome* **8**, 31 (2020).
198. Prudence, S. M. *et al.* Soil, senescence and exudate utilisation: characterisation of the Paragon var. spring bread wheat root microbiome. *Environmental Microbiome* **16**, 12 (2021).
199. Manefield, M., Whiteley, A. S., Griffiths, R. I. & Bailey, M. J. RNA Stable Isotope Probing, a Novel Means of Linking Microbial Community Function to Phylogeny. *Appl Environ Microbiol* **68**, 5367–5373 (2002).
200. Rangel-Castro, J. I. *et al.* Stable isotope probing analysis of the influence of liming on root exudate utilization by soil microorganisms: Stable isotope probing of soil microbial communities. *Environmental Microbiology* **7**, 828–838 (2005).
201. Neufeld, J. D., Dumont, M. G., Vohra, J. & Murrell, J. C. Methodological Considerations for the Use of Stable Isotope Probing in Microbial Ecology. *Microb Ecol* **53**, 435–442 (2007).
202. Nelson, C., Giraldo-Silva, A. & Garcia-Pichel, F. A symbiotic nutrient exchange within the cyanosphere microbiome of the biocrust cyanobacterium, *Microcoleus vaginatus*. *ISME J* **15**, 282–292 (2021).
203. Menyhart, L., Nagy, S. & Lepossa, A. Rapid analysis of photoautotroph microbial communities in soils by flow cytometric barcoding and fingerprinting. *Applied Soil Ecology* **130**, 237–240 (2018).
204. Alam, M. *et al.* Whole-Genome Shotgun Sequence of the Sulfur-Oxidizing Chemoautotroph *Pseudomonas salicylatoxidans* KCT001. *Journal of Bacteriology* **194**, 4743–4744 (2012).
205. Lehtovirta-Morley, L. E. *et al.* Isolation of ' *Candidatus Nitrosocosmicus franklandus*', a novel ureolytic soil archaeal ammonia oxidiser with tolerance to high ammonia concentration. *FEMS Microbiology Ecology* **92**, fiw057 (2016).
206. Bruggeling, C. E. *et al.* Optimized bacterial DNA isolation method for microbiome analysis of human tissues. *MicrobiologyOpen* **10**, (2021).
207. Heravi, F. S., Zakrzewski, M., Vickery, K. & Hu, H. Host DNA depletion efficiency of microbiome DNA enrichment methods in infected tissue samples. *Journal of Microbiological Methods* **170**, 105856 (2020).
208. Ganda, E. *et al.* DNA Extraction and Host Depletion Methods Significantly Impact and Potentially Bias Bacterial Detection in a Biological Fluid. **6**, 19 (2021).
209. Fitzpatrick, C. R. *et al.* Chloroplast sequence variation and the efficacy of peptide nucleic acids for blocking host amplification in plant microbiome studies. *Microbiome* **6**, 144 (2018).
210. Frank, A. C., Guzmán, J. P. S. & Shay, J. E. Transmission of Bacterial Endophytes. *Microorganisms* **10**, 4 (2017).

211. Vujanovic, V. & Germida, J. Seed endosymbiosis: a vital relationship in providing prenatal care to plants. *Can. J. Plant Sci.* CJPS-2016-0261 (2017) doi:10.1139/CJPS-2016-0261.
212. Jiang, L. *et al.* Whole-genome sequence data and analysis of *Saccharibacillus* sp. ATSA2 isolated from Kimchi cabbage seeds. *Data in Brief* **26**, 104465 (2019).
213. Links, M. G. *et al.* Simultaneous profiling of seed-associated bacteria and fungi reveals antagonistic interactions between microorganisms within a shared epiphytic microbiome on *Triticum* and *Brassica* seeds. *New Phytol* **202**, 542–553 (2014).
214. Chimwamurombe, P. M., Grönemeyer, J. L. & Reinhold-Hurek, B. Isolation and characterization of culturable seed-associated bacterial endophytes from gnotobiotically grown Marama bean seedlings. *FEMS Microbiology Ecology* **92**, fiw083 (2016).
215. Hone, H. *et al.* Profiling, isolation and characterisation of beneficial microbes from the seed microbiomes of drought tolerant wheat. *Sci Rep* **11**, 11916 (2021).
216. Rahman, M. M. *et al.* Consistent associations with beneficial bacteria in the seed endosphere of barley (*Hordeum vulgare* L.). *Systematic and Applied Microbiology* **41**, 386–398 (2018).
217. Mahmood, A., Turgay, O. C., Farooq, M. & Hayat, R. Seed biopriming with plant growth promoting rhizobacteria: a review. *FEMS Microbiology Ecology* **92**, fiw112 (2016).
218. Bergna, A. *et al.* Tomato Seeds Preferably Transmit Plant Beneficial Endophytes. *Phytobiomes Journal* **2**, 183–193 (2018).
219. Taffner, J., Bergna, A., Cernava, T. & Berg, G. Tomato-Associated Archaea Show a Cultivar-Specific Rhizosphere Effect but an Unspecific Transmission by Seeds. *Phytobiomes Journal* **4**, 133–141 (2020).
220. Huang, Y., Kuang, Z., Wang, W. & Cao, L. Exploring potential bacterial and fungal biocontrol agents transmitted from seeds to sprouts of wheat. *Biological Control* **98**, 27–33 (2016).
221. Walsh, C. M., Becker-Uncapher, I., Carlson, M. & Fierer, N. Variable influences of soil and seed-associated bacterial communities on the assembly of seedling microbiomes. *ISME J* (2021) doi:10.1038/s41396-021-00967-1.
222. Morales Moreira, Z. P., Helgason, B. L. & Germida, J. J. Crop, genotype, and field environmental conditions shape bacterial and fungal seed epiphytic microbiomes. *Can. J. Microbiol.* **67**, 161–173 (2021).
223. Distelfeld, A., Avni, R. & Fischer, A. M. Senescence, nutrient remobilization, and yield in wheat and barley. *Journal of Experimental Botany* **65**, 3783–3798 (2014).
224. Guiboileau, A., Sormani, R., Meyer, C. & Masclaux-Daubresse, C. Senescence and death of plant organs: Nutrient recycling and developmental regulation. *Comptes Rendus Biologies* **333**, 382–391 (2010).
225. Häffner, E., Konietzki, S. & Diederichsen, E. Keeping Control: The Role of Senescence and Development in Plant Pathogenesis and Defense. *Plants* **4**, 449–488 (2015).
226. Wang, H., Su, J., Zheng, T. & Yang, X. Insights into the role of plant on ammonia-oxidizing bacteria and archaea in the mangrove ecosystem. *J Soils Sediments* **15**, 1212–1223 (2015).
227. Großkopf, R., Janssen, P. H. & Liesack, W. Diversity and Structure of the Methanogenic Community in Anoxic Rice Paddy Soil Microcosms as Examined by Cultivation and Direct 16S rRNA Gene Sequence Retrieval. *Appl. Environ. Microbiol.* **64**, 960–969 (1998).
228. Chen, X.-P., Zhu, Y.-G., Xia, Y., Shen, J.-P. & He, J.-Z. Ammonia-oxidizing archaea: important players in paddy rhizosphere soil? *Environmental Microbiology* **10**, 1978–1987 (2008).
229. Zhang, F. *et al.* Symbiotic archaea in marine sponges show stability and host specificity in community structure and ammonia oxidation functionality. *FEMS Microbiol Ecol* **90**, 699–707 (2014).
230. Probst, A. J., Auerbach, A. K. & Moissl-Eichinger, C. Archaea on Human Skin. *PLoS ONE* **8**, e65388 (2013).
231. Zhang, S. *et al.* Ammonia oxidizers in river sediments of the Qinghai-Tibet Plateau and their adaptations to high-elevation conditions. *Water Research* **173**, 115589 (2020).
232. Lehtovirta, L. E., Prosser, J. I. & Nicol, G. W. Soil pH regulates the abundance and diversity of Group 1.1c Crenarchaeota. *FEMS Microbiol Ecol* **10** (2009).

233. Jung, M.-Y. *et al.* Enrichment and Characterization of an Autotrophic Ammonia-Oxidizing Archaeon of Mesophilic Crenarchaeal Group I.1a from an Agricultural Soil. *Appl. Environ. Microbiol.* **77**, 8635–8647 (2011).
234. Zhalnina, K., de Quadros, P. D., Camargo, F. A. O. & Triplett, E. W. Drivers of archaeal ammonia-oxidizing communities in soil. *Front. Microbio.* **3**, (2012).
235. Baker, G. C., Smith, J. J. & Cowan, D. A. Review and re-analysis of domain-specific 16S primers. *Journal of Microbiological Methods* **55**, 541–555 (2003).
236. Song, G. C. *et al.* Plant growth-promoting archaea trigger induced systemic resistance in *Arabidopsis thaliana* against *Pectobacterium carotovorum* and *Pseudomonas syringae*. *Environ Microbiol* **21**, 940–948 (2019).
237. Zhang, C.-J. *et al.* Diversity, metabolism and cultivation of archaea in mangrove ecosystems. *Mar Life Sci Technol* **3**, 252–262 (2021).
238. Pires, A. C. C. *et al.* Denaturing Gradient Gel Electrophoresis and Barcoded Pyrosequencing Reveal Unprecedented Archaeal Diversity in Mangrove Sediment and Rhizosphere Samples. *Appl. Environ. Microbiol.* **78**, 5520–5528 (2012).
239. Taffner, J. *et al.* What Is the Role of Archaea in Plants? New Insights from the Vegetation of Alpine Bogs. *mSphere* **3**, (2018).
240. Liu, Y., Li, H., Liu, Q. F. & Li, Y. H. Archaeal communities associated with roots of the common reed (*Phragmites australis*) in Beijing Cuihu Wetland. *World J Microbiol Biotechnol* **31**, 823–832 (2015).
241. He, Y. *et al.* Abundance and diversity of ammonia-oxidizing archaea and bacteria in the rhizosphere soil of three plants in the Ebinur Lake wetland. *Can. J. Microbiol.* **63**, 573–582 (2017).
242. Pump, J., Pratscher, J. & Conrad, R. Colonization of rice roots with methanogenic archaea controls photosynthesis-derived methane emission: Photosynthesis-derived methane emission. *Environ Microbiol* **17**, 2254–2260 (2015).
243. Knief, C. *et al.* Metaproteogenomic analysis of microbial communities in the phyllosphere and rhizosphere of rice. *ISME J* **6**, 1378–1390 (2012).
244. Berghuis, B. A. *et al.* Hydrogenotrophic methanogenesis in archaeal phylum Verstraetearchaeota reveals the shared ancestry of all methanogens. *Proc Natl Acad Sci USA* **116**, 5037–5044 (2019).
245. Bridgman, S. D., Cadillo-Quiroz, H., Keller, J. K. & Zhuang, Q. Methane emissions from wetlands: biogeochemical, microbial, and modeling perspectives from local to global scales. *Glob Change Biol* **19**, 1325–1346 (2013).
246. Liu, S., Coyne, M. S., Grove, J. H. & Flythe, M. D. Tillage, not fertilization, dominantly influences ammonia-oxidizing archaea diversity in long-term, continuous maize. *Applied Soil Ecology* **147**, 103384 (2020).
247. Wattenburger, C. J. *et al.* The rhizosphere and cropping system, but not arbuscular mycorrhizae, affect ammonia oxidizing archaea and bacteria abundances in two agricultural soils. *Applied Soil Ecology* **151**, 103540 (2020).
248. Pratscher, J., Dumont, M. G. & Conrad, R. Ammonia oxidation coupled to CO<sub>2</sub> fixation by archaea and bacteria in an agricultural soil. *Proceedings of the National Academy of Sciences* **108**, 4170–4175 (2011).
249. Zhalnina, K. *et al.* Ca. *Nitrososphaera* and *Bradyrhizobium* are inversely correlated and related to agricultural practices in long-term field experiments. *Front. Microbiol.* **4**, (2013).
250. Dignam, B. E. A. *et al.* Effect of land use and soil organic matter quality on the structure and function of microbial communities in pastoral soils: Implications for disease suppression. *PLoS ONE* **13**, e0196581 (2018).
251. Alves, R. J. E., Minh, B. Q., Urich, T., von Haeseler, A. & Schleper, C. Unifying the global phylogeny and environmental distribution of ammonia-oxidising archaea based on amoA genes. *Nat Commun* **9**, 1517 (2018).
252. Rinke, C. *et al.* A standardized archaeal taxonomy for the Genome Taxonomy Database. *Nat Microbiol* **6**, 946–959 (2021).

253. Oliveira, M. N. V. *et al.* Endophytic microbial diversity in coffee cherries of *Coffea arabica* from southeastern Brazil. *Can. J. Microbiol.* **59**, 221–230 (2013).
254. Müller, H. *et al.* Plant genotype-specific archaeal and bacterial endophytes but similar *Bacillus* antagonists colonize Mediterranean olive trees. *Front. Microbiol.* **6**, (2015).
255. Sliwinski, M. K. & Goodman, R. M. Comparison of Crenarchaeal Consortia Inhabiting the Rhizosphere of Diverse Terrestrial Plants with Those in Bulk Soil in Native Environments. *Appl. Environ. Microbiol.* **70**, 1821–1826 (2004).
256. Bomberg, M. & Timonen, S. Effect of Tree Species and Mycorrhizal Colonization on the Archaeal Population of Boreal Forest Rhizospheres. *Appl Environ Microbiol* **75**, 308–315 (2009).
257. Rinta-Kanto, J. M. & Timonen, S. Spatial variations in bacterial and archaeal abundance and community composition in boreal forest pine mycorrhizospheres. *European Journal of Soil Biology* **97**, 103168 (2020).
258. Zhang, M. *et al.* Deciphering the archaeal communities in tree rhizosphere of the Qinghai-Tibetan plateau. *BMC Microbiol* **20**, 235 (2020).
259. Dubey, G., Kollah, B., Gour, V. K., Shukla, A. K. & Mohanty, S. R. Diversity of bacteria and archaea in the rhizosphere of bioenergy crop *Jatropha curcas*. *3 Biotech* **6**, 257 (2016).
260. Chelius, M. K. & Triplett, E. W. The Diversity of Archaea and Bacteria in Association with the Roots of *Zea mays* L. *Microb Ecol* **41**, 252–263 (2001).
261. Fadji, A. E., Ayangbenro, A. S. & Babalola, O. O. Organic Farming Enhances the Diversity and Community Structure of Endophytic Archaea and Fungi in Maize Plant: a Shotgun Approach. *J Soil Sci Plant Nutr* **20**, 2587–2599 (2020).
262. Taffner, J., Cernava, T., Erlacher, A. & Berg, G. Novel insights into plant-associated archaea and their functioning in arugula (*Eruca sativa* Mill.). *Journal of Advanced Research* **19**, 39–48 (2019).
263. Simon, H. M. *et al.* Cultivation of Mesophilic Soil Crenarchaeotes in Enrichment Cultures from Plant Roots. *Appl Environ Microbiol* **71**, 4751–4760 (2005).
264. Kandeler, E. & Gerber, H. Short-term assay of soil urease activity using colorimetric determination of ammonium. *Biol Fert Soils* **6**, (1988).
265. Miranda, K. M., Espey, M. G. & Wink, D. A. A Rapid, Simple Spectrophotometric Method for Simultaneous Detection of Nitrate and Nitrite. **10** (2001).
266. R Core Team. R: A language and environment for statistical computing. (R Foundation for Statistical Computing, 2020).
267. Wickham, H., François, R., Henry, L. & Müller, K. dplyr: A Grammar of Data Manipulation. (2021).
268. Dinno, A. dunn.test: Dunn's Test of Multiple Comparisons Using Rank Sums. (2017).
269. Salles, J. F., De Souza, F. A. & van Elsas, J. D. Molecular Method To Assess the Diversity of *Burkholderia* Species in Environmental Samples. *AEM* **68**, 1595–1603 (2002).
270. Dorn-In, S., Bassitta, R., Schwaiger, K., Bauer, J. & Hölzel, C. S. Specific amplification of bacterial DNA by optimized so-called universal bacterial primers in samples rich of plant DNA. *Journal of Microbiological Methods* **7** (2015).
271. Chemidlin Prévost-Bouré, N. *et al.* Validation and Application of a PCR Primer Set to Quantify Fungal Communities in the Soil Environment by Real-Time Quantitative PCR. *PLoS ONE* **6**, e24166 (2011).
272. Muyzer, G., de Waal, E. C. & Uitterlinden, A. G. Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. *Applied and Environmental Microbiology* **59**, 695–700 (1993).
273. Ochsenreiter, T., Selezi, D., Quaiser, A., Bonch-Osmolovskaya, L. & Schleper, C. Diversity and abundance of Crenarchaeota in terrestrial habitats studied by 16S RNA surveys and real time PCR. *Environ Microbiol* **5**, 787–797 (2003).
274. Gantner, S., Andersson, A. F., Alonso-Sáez, L. & Bertilsson, S. Novel primers for 16S rRNA-based archaeal community analyses in environmental samples. *Journal of Microbiological Methods* **84**, 12–18 (2011).

275. Yu, Y., Lee, C., Kim, J. & Hwang, S. Group-specific primer and probe sets to detect methanogenic communities using quantitative real-time polymerase chain reaction. *Biotechnol. Bioeng.* **89**, 670–679 (2005).
276. Ihrmark, K. *et al.* New primers to amplify the fungal ITS2 region - evaluation by 454-sequencing of artificial and natural communities. *FEMS Microbiol Ecol* **82**, 666–677 (2012).
277. White, T. J., Bruns, T., Lee, S. & Taylor, J. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. in *PCR Protocols* 315–322 (Elsevier, 1990). doi:10.1016/B978-0-12-372180-8.50042-1.
278. Vainio, E. J. & Hantula, J. Direct analysis of wood-inhabiting fungi using denaturing gradient gel electrophoresis of amplified ribosomal DNA. 10.
279. Fredriksson, N. J., Hermansson, M. & Wilén, B.-M. The Choice of PCR Primers Has Great Impact on Assessments of Bacterial Community Diversity and Dynamics in a Wastewater Treatment Plant. *PLOS ONE* **8**, 20 (2013).
280. Rastogi, G. A PCR-based toolbox for the culture-independent quantification of total bacterial abundances in plant environments. *Journal of Microbiological Methods* **6** (2010).
281. Findley, F. *et al.* Intramural Sequencing Center Comparative Sequencing Program *et al.* Topographic diversity of fungal and bacterial communities in human skin. *Nature* **498**, 367–370 (2013).
282. Amaral-Zettler, L. A., McCliment, E. A., Ducklow, H. W. & Huse, S. M. A Method for Studying Protistan Diversity Using Massively Parallel Sequencing of V9 Hypervariable Regions of Small-Subunit Ribosomal RNA Genes. *PLoS ONE* **4**, e6372 (2009).
283. Wang, Y., Tian, R. M., Gao, Z. M., Bougouffa, S. & Qian, P.-Y. Optimal Eukaryotic 18S and Universal 16S/18S Ribosomal RNA Primers and Their Application in a Study of Symbiosis. *PLoS ONE* **9**, e90053 (2014).
284. Takai, K. & Horikoshi, K. Rapid Detection and Quantification of Members of the Archaeal Community by Quantitative PCR Using Fluorogenic Probes. *Appl. Environ. Microbiol.* **66**, 5066–5072 (2000).
285. Tourna, M., Freitag, T. E., Nicol, G. W. & Prosser, J. I. Growth, activity and temperature responses of ammonia-oxidizing archaea and bacteria in soil microcosms. *Environmental Microbiology* **9** (2008).
286. Heuer, H., Krsek, M., Baker, P., Smalla, K. & Wellington, E. M. Analysis of actinomycete communities by specific amplification of genes encoding 16S rRNA and gel-electrophoretic separation in denaturing gradients. *Applied and environmental microbiology* **63**, 3233–3241 (1997).
287. Freeman, J. & Ward, E. *Gaeumannomyces graminis*, the take-all fungus and its relatives. *Mol Plant Pathol* **5**, 235–252 (2004).
288. Bolyen, E. *et al.* Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat Biotechnol* **37**, 852–857 (2019).
289. Callahan, B. J. *et al.* DADA2: High-resolution sample inference from Illumina amplicon data. *Nat Methods* **13**, 581–583 (2016).
290. Glöckner, F. O. *et al.* 25 years of serving the community with ribosomal RNA gene reference databases and tools. *Journal of Biotechnology* **261**, 169–176 (2017).
291. Nilsson, R. H. *et al.* The UNITE database for molecular identification of fungi: handling dark taxa and parallel taxonomic classifications. *Nucleic Acids Research* **47**, D259–D264 (2019).
292. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic Local Alignment Search Tool. 8.
293. Jari, O. *et al.* vegan: Community Ecology Package. R package version 2.5-6. (2019).
294. Clarke, K. R. Non-parametric multivariate analyses of changes in community structure. *Austral Ecol* **18**, 117–143 (1993).
295. McMurdie, P. J. & Holmes, S. phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLoS ONE* **8**, e61217 (2013).
296. Ssekagiri, A. microbiomeSeq: Microbial community analysis in an environmental context.. R package version 0.1. (2020).

297. Lin, H. & Peddada, S. D. Analysis of compositions of microbiomes with bias correction. *Nat Commun* **11**, 3514 (2020).
298. Lahti, L. & Shetty, S. microbiome R package. (2019).
299. Pagès, H., Aboyoun, P., Gentleman, R. & DebRoy, S. Biostrings: Efficient manipulation of biological strings. *R package version 2.60.1* (2021).
300. Morgan, M. *et al.* ShortRead: a Bioconductor package for input, quality assessment and exploration of high-throughput sequence data. *Bioinformatics* **25**, 2607–2608 (2009).
301. Wickham, H. ggplot2: Elegant Graphics for Data Analysis. *Springer-Verlag New York* (2016).
302. Wickham, H. Reshaping data with the reshape package. *Journal of Statistical Software* **21**, (2007).
303. Baptiste, A. gridExtra: Miscellaneous Functions for 'Grid' Graphics. *R package version 2.0.0* (2015).
304. Parks, D. H. *et al.* A complete domain-to-species taxonomy for Bacteria and Archaea. *Nat Biotechnol* **38**, 1079–1086 (2020).
305. Madeira, F. *et al.* The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Research* **47**, W636–W641 (2019).
306. Venturi, V. & Keel, C. Signaling in the Rhizosphere. *Trends in Plant Science* **21**, 187–198 (2016).
307. Lofgren, L. A. *et al.* Genome-based estimates of fungal rDNA copy number variation across phylogenetic scales and ecological lifestyles. *Mol Ecol* **28**, 721–730 (2019).
308. Sun, D.-L., Jiang, X., Wu, Q. L. & Zhou, N.-Y. Intragenomic Heterogeneity of 16S rRNA Genes Causes Overestimation of Prokaryotic Diversity. *Appl. Environ. Microbiol.* **79**, 5962–5969 (2013).
309. Jung, M.-Y. *et al.* A hydrophobic ammonia-oxidizing archaeon of the *Nitrosocosmicus* clade isolated from coal tar-contaminated sediment: Ecophysiological properties of a hydrophobic AOA. *Environmental Microbiology Reports* **8**, 983–992 (2016).
310. Newitt, J. T., Prudence, S. M. M., Hutchings, M. I. & Worsley, S. F. Biocontrol of Cereal Crop Diseases Using Streptomyces. *Pathogens* **8**, 78 (2019).
311. Lambert, D. H. *Streptomyces scabies* sp. nov., nom. rev. *International Journal of Systematic Bacteriology* **39**, 6 (1989).
312. Chater, K. F., Biró, S., Lee, K. J., Palmer, T. & Schrempf, H. The complex extracellular biology of *Streptomyces*. *FEMS Microbiol Rev* **34**, 171–198 (2010).
313. Eberl, L. & Vandamme, P. Members of the genus *Burkholderia*: good and bad guys. *F1000Res* **5**, 1007 (2016).
314. Braun, S., Gevens, A., Charkowski, A., Allen, C. & Jansky, S. Potato Common Scab: a Review of the Causal Pathogens, Management Practices, Varietal Resistance Screening Methods, and Host Resistance. *Am. J. Potato Res.* **94**, 283–296 (2017).
315. Thell, A. *et al.* A review of the lichen family *Parmeliaceae* - history, phylogeny and current taxonomy. *Nordic Journal of Botany* **30**, 641–664 (2012).
316. Cao, Q. *et al.* Effects of Rare Microbiome Taxa Filtering on Statistical Analysis. *Front. Microbiol.* **11**, 607325 (2021).
317. Harman, G. E., Howell, C. R., Viterbo, A., Chet, I. & Lorito, M. Trichoderma species — opportunistic, avirulent plant symbionts. *Nat Rev Microbiol* **2**, 43–56 (2004).
318. McMillan, V. E., Gutteridge, R. J. & Hammond-Kosack, K. E. Identifying variation in resistance to the take-all fungus, *Gaeumannomyces graminis* var. *tritici*, between different ancestral and modern wheat species. *12* (2014).
319. Guo, L. D., Hyde, K. D. & Liew, E. C. Y. Identification of endophytic fungi from *Livistona chinensis* based on morphology and rDNA sequences. *New Phytologist* **147**, 617–630 (2000).

320. Daranagama, D. A. *et al.* *Tristratiperidium microsporium* gen. et sp. nov. (Xylariales) on dead leaves of *Arundo plinii*. *Mycol Progress* **15**, 8 (2016).
321. Kumar, N. & Iyer-Pascuzzi, A. S. Shedding the Last Layer: Mechanisms of Root Cap Cell Release. *Plants* **9**, 308 (2020).
322. Tran, T. M., MacIntyre, A., Hawes, M. & Allen, C. Escaping Underground Nets: Extracellular DNases Degrade Plant Extracellular Traps and Contribute to Virulence of the Plant Pathogenic Bacterium *Ralstonia solanacearum*. *PLoS Pathog* **12**, e1005686 (2016).
323. Kumpf, R. P. & Nowack, M. K. The root cap: a short story of life and death. *EXBOTJ* **66**, 5651–5662 (2015).
324. Carrias, J.-F. *et al.* Resource availability drives bacterial succession during leaf-litter decomposition in a bromeliad ecosystem. *FEMS Microbiology Ecology* **96**, fiae045 (2020).
325. Purahong, W. *et al.* Life in leaf litter: novel insights into community dynamics of bacteria and fungi during litter decomposition. *Mol Ecol* **25**, 4059–4074 (2016).
326. de Melo, R. R. *et al.* Unraveling the cellulolytic and hemicellulolytic potential of two novel *Streptomyces* strains. *Ann Microbiol* **68**, 677–688 (2018).
327. Fernández, F. A. & Huhndorf, S. M. New species of *Chaetosphaeria*, *Melanopsammella* and *Tainosphaeria* gen. nov. from the Americas. *Fungal Diversity* **43**.
328. Zhang, Y., Crous, P. W., Schoch, C. L. & Hyde, K. D. Pleosporales. *Fungal Diversity* **53**, 1–221 (2012).
329. Oliver, R. *et al.* Absence of detectable yield penalty associated with insensitivity to Pleosporales necrotrophic effectors in wheat grown in the West Australian wheat belt. *Plant Pathol* **63**, 1027–1032 (2014).
330. Classen, A. T. *et al.* Direct and indirect effects of climate change on soil microbial and soil microbial-plant interactions: What lies ahead? *Ecosphere* **6**, art130 (2015).
331. Réblová, M., Gams, W. & Seifert, K. A. *Monilochaetes* and allied genera of the *Glomerellales*, and a reconsideration of families in the Microascales. *Studies in Mycology* **68**, 163–191 (2011).
332. Chen, H., Ma, Y., Zhang, W. F., Ma, T. & Wu, H. X. Molecular phylogeny of *Colletotrichum* (Sordariomycetes: *Glomerellaceae*) inferred from multiple gene sequences. *Genet. Mol. Res.* **14**, 13649–13662 (2015).
333. Ghosh, R., Bhadra, S., Bandyopadhyay, M. Morphological and molecular characterization of *Colletotrichum capsici* causing leaf-spot of soybean. *Trop. Plant Res.* **3**, 481–490 (2016).
334. Zhang, N. *et al.* An overview of the systematics of the Sordariomycetes based on a four-gene phylogeny. **13**.
335. Kowal, J., Pressel, S., Duckett, J. G., Bidartondo, M. I. & Field, K. J. From rhizoids to roots? Experimental evidence of mutualism between liverworts and ascomycete fungi. *Annals of Botany* **121**, 221–227 (2018).
336. Midgley, D. J., Greenfield, P., Bissett, A. & Tran-Dinh, N. First evidence of *Pezoloma ericae* in Australia: using the Biomes of Australia Soil Environments (BASE) to explore the Australian phylogeography of known ericoid mycorrhizal and root-associated fungi. *Mycorrhiza* **27**, 587–594 (2017).
337. Hambleton, S. & Sigler, L. *Meliniomyces*, a new anamorph genus for root-associated fungi with phylogenetic affinities to *Rhizoscyphus ericae* ( $\equiv$  *Hymenoscyphus ericae*), *Leotiomyces*. *Studies in Mycology* **53**, 1–27 (2005).
338. Fradgley, N. *et al.* A large-scale pedigree resource of wheat reveals evidence for adaptation and selection by breeders. *PLoS Biol* **17**, e3000071 (2019).
339. Nelson, E. B. The seed microbiome: Origins, interactions, and impacts. *Plant Soil* **422**, 7–34 (2018).
340. Liu, Y. *et al.* Tobacco bacterial wilt can be biologically controlled by the application of antagonistic strains in combination with organic fertilizer. *Biol Fertil Soils* **18** (2013).

341. Ding, C., Shen, Q., Zhang, R. & Chen, W. Evaluation of rhizosphere bacteria and derived bio-organic fertilizers as potential biocontrol agents against bacterial wilt (*Ralstonia solanacearum*) of potato. *Plant Soil* **366**, 453–466 (2013).
342. Edwards, J. *et al.* Structure, variation, and assembly of the root-associated microbiomes of rice. *plant biology* **10**.
343. Chaparro, J. M. *et al.* Root Exudation of Phytochemicals in Arabidopsis Follows Specific Patterns That Are Developmentally Programmed and Correlate with Soil Microbial Functions. *PLoS ONE* **8**, e55731 (2013).
344. Bulgarelli, D., Schlaeppli, K., Spaepen, S., van Themaat, E. V. L. & Schulze-Lefert, P. Structure and Functions of the Bacterial Microbiota of Plants. *Annu. Rev. Plant Biol.* **64**, 807–838 (2013).
345. Kulik, T., Treder, K. & Zaluski, D. Quantification of *Alternaria*, *Cladosporium*, *Fusarium* and *Penicillium verrucosum* in Conventional and Organic Grains by qPCR. *J Phytopathol* **163**, 522–528 (2015).
346. Perelló, A. E., Sisterna, M. N. & Moreno, M. V. Occurrence of *Cladosporium herbarum* on wheat leaves (*Triticum aestivum*) in Argentina. *Austral. Plant Pathol.* **32**, 327 (2003).
347. Marshall, R. *et al.* Analysis of Two in Planta Expressed LysM Effector Homologs from the Fungus *Mycosphaerella graminicola* Reveals Novel Functional Properties and Varying Contributions to Virulence on Wheat. *Plant Physiology* **156**, 756–769 (2011).
348. Jeon, S. J., Nguyen, T. T. T. & Lee, H. B. Phylogenetic Status of an Unrecorded Species of *Curvularia*, *C. spicifera*, Based on Current Classification System of *Curvularia* and *Bipolaris* Group Using Multi Loci. *Mycobiology* **43**, 210–217 (2015).
349. Kim, D.-R. *et al.* A mutualistic interaction between *Streptomyces* bacteria, strawberry plants and pollinating bees. *Nat Commun* **10**, 4802 (2019).
350. Neal, A. L., Ahmad, S., Gordon-Weeks, R. & Ton, J. Benzoxazinoids in Root Exudates of Maize Attract *Pseudomonas putida* to the Rhizosphere. *PLoS ONE* **7**, 10 (2012).
351. Arunkumar, N., Rakesh, S., Rajaram, K., Kumar, N. R. & Durairajan, S. S. K. Anthosphere Microbiome and Their Associated Interactions at the Aromatic Interface. in *Plant Microbe Interface*. 309–324. doi:10.1007/978-3-030-19831-2\_14.
352. Droby, S. & Sionov, E. Shifts in the Composition of the Microbiota of Stored Wheat Grains in Response to Fumigation. *Frontiers in Microbiology* **10**, 13 (2019).
353. Cretoi, M. S., Kielak, A. M., Schluter, A. & van Elsas, J. D. Bacterial communities in chitin-amended soil as revealed by 16S rRNA gene based pyrosequencing. *Soil Biology and Biochemistry* **76**, 5–11 (2014).
354. Ofek, M., Hadar, Y. & Minz, D. Ecology of Root Colonizing *Massilia* (*Oxalobacteraceae*). *PLoS ONE* **7**, e40117 (2012).
355. Araujo, R., Dunlap, C., Barnett, S. & Franco, C. M. M. Decoding Wheat Endosphere–Rhizosphere Microbiomes in *Rhizoctonia solani*–Infested Soils Challenged by *Streptomyces* Biocontrol Agents. *Front. Plant Sci.* **10**, 1038 (2019).
356. Du, C. *Massilia cellulositytica* sp. nov., a novel cellulose-degrading bacterium isolated from rhizosphere soil of rice (*Oryza sativa* L.) and its whole genome analysis. *Antonie van Leeuwenhoek* **12**.
357. Anderson, M. & Habiger, J. Characterization and Identification of Productivity-Associated Rhizobacteria in Wheat. *Appl. Environ. Microbiol.* **78**, 4434–4446 (2012).
358. Fang, K. *et al.* Growth-promoting characteristics of potential nitrogen-fixing bacteria in the root of an invasive plant *Ageratina adenophora*. *PeerJ* **7**, e7099 (2019).
359. Aranda, S., Montes-Borrego, M. & Landa, B. B. Purple-Pigmented Violacein-Producing *Duganella* spp. Inhabit the Rhizosphere of Wild and Cultivated Olives in Southern Spain. *Microb Ecol* **62**, 446–459 (2011).
360. Cardinale, M., Suarez, C., Steffens, D., Ratering, S. & Schnell, S. Effect of Different Soil Phosphate Sources on the Active Bacterial Microbiota Is Greater in the Rhizosphere than in the Endorhiza of Barley (*Hordeum vulgare* L.). *Microb Ecol* **77**, 689–700 (2019).

361. Miyajima, K., Tanaka, F., Takeuchi, T. & Kuninaga, S. *Streptomyces turgidiscabies* sp. nov. *International Journal of Systematic Bacteriology* **48**, 495–502 (1998).
362. Zhang, Y. *et al.* Antifungal Activities of Metabolites Produced by a Termite-Associated *Streptomyces canus* BYB02. *J. Agric. Food Chem.* **61**, 1521–1524 (2013).
363. Wu, F. *et al.* *Streptomyces canus* GLY-P2 degrades ferulic and p-hydroxybenzoic acids in soil and affects cucumber antioxidant enzyme activity and rhizosphere bacterial community. *Plant Soil* **436**, 71–89 (2019).
364. Poovarasan, S., Mohandas, S., Paneerselvam, P., Saritha, B. & Ajay, K. M. Mycorrhizae colonizing actinomycetes promote plant growth and control bacterial blight disease of pomegranate (*Punica granatum* L. cv Bhagwa). *Crop Protection* **53**, 175–181 (2013).
365. Wen, C. *et al.* *Streptomyces scabiei* subsp. *xuchangensis*, A novel streptomycete isolate for staurosporine production and a wheat take-all control agent. *Int J of Micr Res* **4**, 282–289 (2012).
366. Gao, B., Paramanathan, R. & Gupta, R. S. Signature proteins that are distinctive characteristics of Actinobacteria and their subgroups. *Antonie Van Leeuwenhoek* **90**, 69–91 (2006).
367. Antony-Babu, S. *et al.* Multiple *Streptomyces* species with distinct secondary metabolomes have identical 16S rRNA gene sequences. *Sci Rep* **7**, 11089 (2017).
368. Ogier, J.-C., Pagès, S., Galan, M., Barret, M. & Gaudriault, S. *rpoB*, a promising marker for analyzing the diversity of bacterial communities by amplicon sequencing. *BMC Microbiol* **19**, 171 (2019).
369. Guo, Y., Zheng, W., Rong, X. & Huang, Y. A multilocus phylogeny of the *Streptomyces griseus* 16S rRNA gene clade: use of multilocus sequence analysis for streptomycete systematics. *International Journal Of Systematic And Evolutionary Microbiology* **58**, 149–159 (2008).
370. Müller, H. *et al.* Complete Genome Sequence of the Sugar Beet Endophyte *Pseudomonas poae* RE\*1-1-14, a Disease-Suppressive Bacterium. *Genome Announc* **1**, (2013).
371. Xia, Y. *et al.* Improved Draft Genome Sequence of *Pseudomonas poae* A2-S9, a Strain with Plant Growth-Promoting Activity. *Microbiol Resour Announc* **8**, (2019).
372. Ren, Y. *et al.* Isolation and characterization of a *Pseudomonas poae* JSU-Y1 with patulin degradation ability and biocontrol potential against *Penicillium expansum*. *Toxicon* **195**, 1–6 (2021).
373. Irshad, U. & Yergeau, E. Bacterial Subspecies Variation and Nematode Grazing Change P Dynamics in the Wheat Rhizosphere. *Front. Microbiol.* **9**, 1990 (2018).
374. Verma, P. *et al.* Appraisal of diversity and functional attributes of thermotolerant wheat associated bacteria from the peninsular zone of India. *Saudi Journal of Biological Sciences* **26**, 1882–1895 (2019).
375. Österman, J., Mousavi, S. A., Koskinen, P., Paulin, L. & Lindström, K. Genomic features separating ten strains of *Neorhizobium galegae* with different symbiotic phenotypes. *BMC Genomics* **16**, 348 (2015).
376. Gage, D. J. Infection and Invasion of Roots by Symbiotic, Nitrogen-Fixing *Rhizobia* during Nodulation of Temperate Legumes. *Microbiol. Mol. Biol. Rev.* **68**, 21 (2004).
377. Yanni, Y. G. *et al.* Assessment of the natural endophytic association between *Rhizobium* and wheat and its ability to increase wheat production in the Nile delta. *Plant Soil* **407**, 367–383 (2016).
378. Paul, S. *Rhizobial* interaction with wheat under water stress condition. *Indian Journal of Plant Physiology* **2**, 131–134 (1997).
379. Mantelin, S. *et al.* Emended description of the genus *Phyllobacterium* and description of four novel species associated with plant roots: *Phyllobacterium bourgognense* sp. nov., *Phyllobacterium ifriqiyense* sp. nov., *Phyllobacterium leguminum* sp. nov. and *Phyllobacterium brassicacearum* sp. nov. *International Journal of Systematic and Evolutionary Microbiology* **56**, 827–839 (2006).
380. Chen, Y. P. *et al.* Phosphate solubilizing bacteria from subtropical soil and their tricalcium phosphate solubilizing abilities. *Applied Soil Ecology* **34**, 33–41 (2006).
381. Breitzkreuz, C., Buscot, F., Tarkka, M. & Reitz, T. Shifts Between and Among Populations of Wheat Rhizosphere *Pseudomonas*, *Streptomyces* and *Phyllobacterium* Suggest Consistent Phosphate Mobilization at Different Wheat Growth Stages Under Abiotic Stress. *Front. Microbiol.* **10**, 3109 (2020).

382. Cheng, M. *et al.* Genetic Transformation of Wheat Mediated by *Agrobacterium tumefaciens*. *Plant Physiology* **115**, 971–980 (1997).
383. Yan, Z.-F. *et al.* *Niastella hibisci* sp. nov., isolated from rhizosphere soil of mugunghwa, the Korean national flower. *International Journal of Systematic and Evolutionary Microbiology* **66**, 5218–5222 (2016).
384. Barajas, H. R. *et al.* Testing the Two-Step Model of Plant Root Microbiome Acquisition Under Multiple Plant Species and Soil Sources. *Front. Microbiol.* **11**, 542742 (2020).
385. Lian, T.-X. *et al.* Bacterial communities incorporating plant-derived carbon in the soybean rhizosphere in Mollisols that differ in soil organic carbon content. *Applied Soil Ecology* **119**, 375–383 (2017).
386. Gao, Y., Lu, Y., Lin, W., Tian, J. & Cai, K. Biochar Suppresses Bacterial Wilt of Tomato by Improving Soil Chemical Properties and Shifting Soil Microbial Community. *Microorganisms* **7**, 676 (2019).
387. Kolton, M. *et al.* Impact of Biochar Application to Soil on the Root-Associated Bacterial Community Structure of Fully Developed Greenhouse Pepper Plants. *Appl. Environ. Microbiol.* **77**, 4924–4930 (2011).
388. Yin, C. *et al.* Rhizosphere community selection reveals bacteria associated with reduced root disease. *Microbiome* **9**, 86 (2021).
389. Han, S.-I., Lee, Y.-R., Kim, J.-O. & Whang, K.-S. *Terrimonas rhizosphaerae* sp. nov., isolated from ginseng rhizosphere soil. *International Journal of Systematic and Evolutionary Microbiology* **67**, 391–395 (2017).
390. Kim, S.-J. *et al.* *Terrimonas terrae* sp. nov., isolated from the rhizosphere of a tomato plant. *International Journal of Systematic and Evolutionary Microbiology* **67**, 3105–3110 (2017).
391. Haichar, F. el Z. *et al.* Plant host habitat and root exudates shape soil bacterial community structure. *ISME J* **2**, 1221–1230 (2008).
392. Kong, Y. *et al.* DNA Stable-Isotope Probing Delineates Carbon Flows from Rice Residues into Soil Microbial Communities Depending on Fertilization. *Appl Environ Microbiol* **86**, e02151-19, /aem/86/7/AEM.02151-19.atom (2020).
393. Vandenkoornhuysen, P. *et al.* Active root-inhabiting microbes identified by rapid incorporation of plant-derived carbon into RNA. *Proceedings of the National Academy of Sciences* **104**, 16970–16975 (2007).
394. Hünninghaus, M. *et al.* Disentangling carbon flow across microbial kingdoms in the rhizosphere of maize. *Soil Biology and Biochemistry* **134**, 122–130 (2019).
395. Wang, J., Chapman, S. J. & Yao, H. Incorporation of <sup>13</sup>C-labelled rice rhizodeposition into soil microbial communities under different fertilizer applications. *Applied Soil Ecology* **101**, 11–19 (2016).
396. Hannula, S. E. *et al.* Shifts in rhizosphere fungal community during secondary succession following abandonment from agriculture. *ISME J* **11**, 2294–2304 (2017).
397. Santander, C. *et al.* Arbuscular mycorrhizal fungal abundance in elevation belts of the hyperarid Atacama Desert. *Fungal Ecology* **51**, 101060 (2021).
398. Martín-Robles, N. *et al.* Impacts of domestication on the arbuscular mycorrhizal symbiosis of 27 crop species. *New Phytol* **218**, 322–334 (2018).
399. Luo, X. *et al.* Arbuscular mycorrhizal fungal communities of topsoil and subsoil of an annual maize-wheat rotation after 15-years of differential mineral and organic fertilization. *Agriculture, Ecosystems & Environment* **315**, 107442 (2021).
400. Mercado-Blanco, J. & Bakker, P. A. H. M. Interactions between plants and beneficial *Pseudomonas* spp.: exploiting bacterial traits for crop protection. *Antonie van Leeuwenhoek* **92**, 367–389 (2007).
401. Rico, A., McCraw, S. L. & Preston, G. M. The metabolic interface between *Pseudomonas syringae* and plant cells. *Current Opinion in Microbiology* **14**, 31–38 (2011).
402. Bais, H. P., Weir, T. L., Perry, L. G., Gilroy, S. & Vivanco, J. M. The role of root exudates in rhizosphere interactions with plants and other organisms. *Annu. Rev. Plant Biol.* **57**, 233–266 (2006).

403. Ortet, P. *et al.* Complete Genome Sequence of a Beneficial Plant Root-Associated Bacterium, *Pseudomonas brassicacearum*. *Journal of Bacteriology* **193**, 12. (2011).
404. Silby, M. W. *et al.* Genomic and genetic analyses of diversity and plant interactions of *Pseudomonas fluorescens*. *Genome Biol* **10**, R51 (2009).
405. Yaghoubi Khanghahi, M., Strafella, S., Allegretta, I. & Crecchio, C. Isolation of Bacteria with Potential Plant-Promoting Traits and Optimization of Their Growth Conditions. *Curr Microbiol* **78**, 464–478 (2021).
406. Sahin, N. Oxalotrophic bacteria. *Research in Microbiology* **154**, 399–407 (2003).
407. Kost, T., Stopnisek, N., Agnoli, K., Eberl, L. & Weiskopf, L. Oxalotrophy, a widespread trait of plant-associated *Burkholderia* species, is involved in successful root colonization of lupin and maize by *Burkholderia phytofirmans*. *Front. Microbiol.* **4**, (2014).
408. Krishnan, R. *et al.* *Arthrobacter pokkali* sp nov, a Novel Plant Associated Actinobacterium with Plant Beneficial Properties, Isolated from Saline Tolerant Pokkali Rice, Kerala, India. *PLoS ONE* **11**, e0150322 (2016).
409. Manzanera, M., Narváez-Reinaldo, J. J., García-Fontana, C., Vílchez, J. I. & González-López, J. Genome Sequence of *Arthrobacter koreensis* 5J12A, a Plant Growth-Promoting and Desiccation-Tolerant Strain. *Genome Announc* **3**, (2015).
410. Shimasaki, T. *et al.* Tobacco Root Endophytic *Arthrobacter* Harbors Genomic Features Enabling the Catabolism of Host-Specific Plant Specialized Metabolites. **12**, 16 (2021).
411. Patel, J. K. & Archana, G. Diverse culturable diazotrophic endophytic bacteria from *Poaceae* plants show cross-colonization and plant growth promotion in wheat. *Plant Soil* **417**, 99–116 (2017).
412. Podar, M., Turner, J., Burdick, L. H. & Pelletier, D. A. Complete Genome Sequence of the Novel *Roseimicrobium* sp. Strain ORNL1, a *Verrucomicrobium* Isolated from the *Populus deltoides* Rhizosphere. *Microbiol Resour Announc* **9**, (2020).
413. van Passel, M. W. J. *et al.* Genome Sequence of the *Verrucomicrobium Opitutus terrae* PB90-1, an Abundant Inhabitant of Rice Paddy Soil Ecosystems. *J Bacteriol* **193**, 2367–2368 (2011).
414. Garcia-Fraile, P., Velazquez, E., Mateos, P. F., Martinez-Molina, E. & Rivas, R. *Cohnella phaseoli* sp. nov., isolated from root nodules of *Phaseolus coccineus* in Spain, and emended description of the genus *Cohnella*. *International Journal Of Systematic And Evolutionary Microbiology* **58**, 1855–1859 (2008).
415. Flores-Félix, J. D. *et al.* *Cohnella lupini* sp. nov., an endophytic bacterium isolated from root nodules of *Lupinus albus*. *International Journal of Systematic and Evolutionary Microbiology* **64**, 83–87 (2014).
416. Rivas, R. *et al.* *Saccharibacillus sacchari* gen. nov., sp. nov., isolated from sugar cane. *International Journal Of Systematic And Evolutionary Microbiology* **58**, 1850–1854 (2008).
417. Kämpfer, P., Busse, H.-J., Kleinhagauer, T., McInroy, J. A. & Glaeser, S. P. *Saccharibacillus endophyticus* sp. nov., an endophyte of cotton. *International Journal of Systematic and Evolutionary Microbiology* **66**, 5134–5139 (2016).
418. Besaury, L. & Remond, C. Draft Genome Sequence of *Saccharibacillus* sp. Strain WB 17, Isolated from Wheat Phyllosphere. *Microbiol Resour Announc* **9**, MRA.01201-19, e01201-19 (2020).
419. Darji, H. *et al.* Polyphasic characterization of and genomic insights into a haloalkali-tolerant *Saccharibacillus alkalitolerans* sp. nov., that produces three cellulase isozymes and several antimicrobial compounds. *Antonie van Leeuwenhoek* (2021) doi:10.1007/s10482-021-01575-x.
420. Rakotoarivonina, H., Hermant, B., Monthe, N. & Rémond, C. The hemicellulolytic enzyme arsenal of *Thermobacillus xylanilyticus* depends on the composition of biomass used for growth. *Microb Cell Fact* **11**, 159 (2012).
421. Liu, X., Li, Q., Li, Y., Guan, G. & Chen, S. *Paenibacillus* strains with nitrogen fixation and multiple beneficial properties for promoting plant growth. *PeerJ* **7**, e7445 (2019).
422. Kim, Y.-J. *et al.* Complete genome sequence of *Paenibacillus yonginensis* DCY84T, a novel plant symbiont that promotes growth via induced systemic resistance. *Stand in Genomic Sci* **12**, 63 (2017).

423. Rybakova, D. *et al.* Endophytes-assisted biocontrol: novel insights in ecology and the mode of action of *Paenibacillus*. *Plant Soil* **405**, 125–140 (2016).
424. McBride, M. J., Liu, W., Lu, X., Zhu, Y. & Zhang, W. The Family *Cytophagaceae*. in *The Prokaryotes*. 577–593. doi:10.1007/978-3-642-38954-2\_382.
425. Zhu, Y. & McBride, M. J. The unusual cellulose utilization system of the aerobic soil bacterium *Cytophaga hutchinsonii*. *Appl Microbiol Biotechnol* **101**, 7113–7127 (2017).
426. Matheron, C., Delort, A.-M., Gaudet, G. & Forano, E. In vivo <sup>13</sup>C NMR study of glucose and cellobiose metabolism by four cellulolytic strains of the genus *Fibrobacter*. 12.
427. Prosser, J. I. & Tough, A. J. Growth mechanisms and growth kinetics of filamentous microorganisms. *Biotechnology* **10**, 253–274 (1991).
428. Szafran, M. J., Jakimowicz, D. & Elliot, M. A. Compaction and control—the role of chromosome-organizing proteins in *Streptomyces*. *FEMS Microbiology Reviews* **44**, 725–739 (2020).
429. Maheshwari, R. Nuclear behavior in fungal hyphae. *FEMS Microbiology Letters* **249**, 7–14 (2005).
430. Parks, D. H. *et al.* A complete domain-to-species taxonomy for Bacteria and Archaea. *Nat Biotechnol* **38**, 1079–1086 (2020).
431. Castanheira, N. *et al.* Plant growth-promoting *Burkholderia* species isolated from annual ryegrass in Portuguese soils. *J Appl Microbiol* **120**, 724–739 (2016).
432. Flórez, L. V. *et al.* Antibiotic-producing symbionts dynamically transition between plant pathogenicity and insect-defensive mutualism. *Nat Commun* **8**, 15172 (2017).
433. Paungfoo-Lonhienne, C. *et al.* A new species of *Burkholderia* isolated from sugarcane roots promotes plant growth. *Microbial Biotechnology* **7**, 142–154 (2014).
434. Palaniappan, P., Chauhan, P. S., Saravanan, V. S., Anandham, R. & Sa, T. Isolation and characterization of plant growth promoting endophytic bacterial isolates from root nodule of *Lespedeza* sp. *Biol Fertil Soils* **46**, 807–816 (2010).
435. Smith, J. L., Goldberg, J. M. & Grossman, A. D. Complete Genome Sequences of *Bacillus subtilis* subsp. *subtilis* Laboratory Strains JH642 (AG174) and AG1839. *Genome Announcements* **2**, e00663-14, 2/4/e00663-14 (2014).
436. Jones, S. E. *et al.* *Streptomyces* Volatile Compounds Influence Exploration and Microbial Community Dynamics by Altering Iron Availability. **10**, 18 (2019).
437. Salanoubat, M. *et al.* Genome sequence of the plant pathogen. **415**, 6 (2002).
438. Rix, G. D., Todd, J. D., Neal, A. L. & Brearley, C. A. Improved sensitivity, accuracy and prediction provided by a high-performance liquid chromatography screen for the isolation of phytase-harboring organisms from environmental samples. *Microb. Biotechnol.* **14**, 1409–1421 (2021).

## Supplementary Figures and Tables

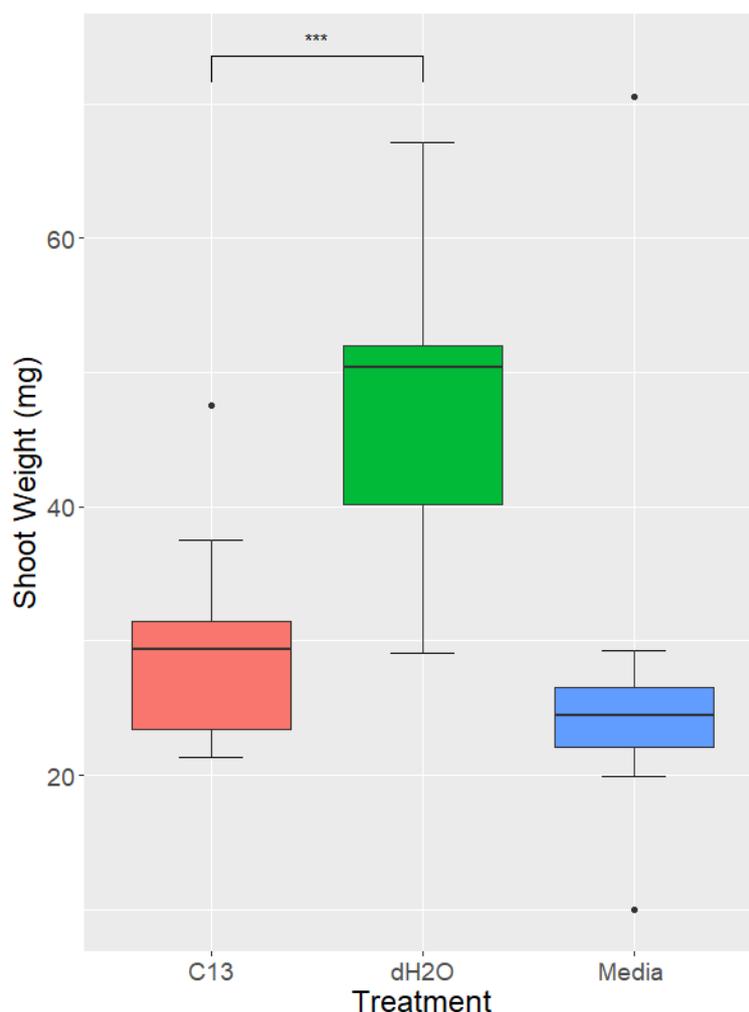


Figure S.1. Experiment to test for plant growth promotion by *Nitrosocosmicus franklandus* C13 of *Arabidopsis thaliana* Col-0. Boxes show *A. thaliana* shoot fresh weights (mg) after treatment with *N. franklandus* C13 suspended in AOA growth medium (red), an H<sub>2</sub>O negative control (green), or an uninoculated ammonia oxidising archaea (AOA) growth medium control (blue). Kruskal-Wallis rank sum test showed a significant effect of treatment on shoot fresh weights (chi-squared = 15.7,  $p < 0.001$ ). While treatment with C13 significantly reduced shoot fresh weight compared to the H<sub>2</sub>O control (Dunns test,  $p < 0.01$ ), a similar significant reduction was observed when comparing plants treated with uninoculated AOA growth media to the H<sub>2</sub>O control ( $p < 0.001$ ). Plants were cultivated in Levingtons F2 compost, N=12 per treatment. This data demonstrated that the AOA growth medium had an inhibitory effect on plant growth.

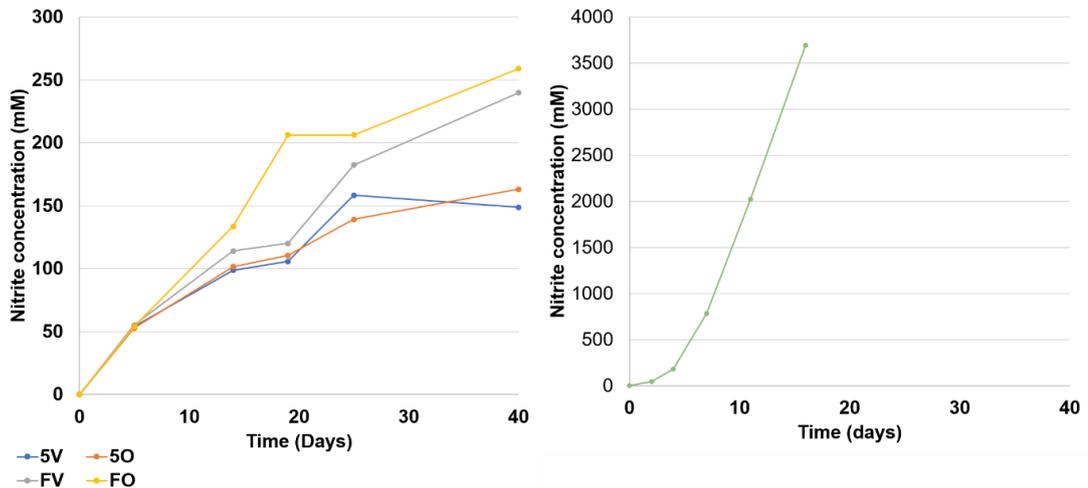


Figure S.2. Experiment to test for the activity of ammonia oxidising archaeon (AOA) *Nitrosocosmicus franklandus* C13 in MSK plant growth medium. Left shows the nitrite concentration (mM), used as a proxy for ammonia oxidation, in four enrichment culture conditions over 40 days, **5V** half strength MSK medium supplemented with AOA medium vitamins (blue), **5O** half strength MSK medium (orange), **FV** full strength MSK medium supplemented with AOA medium vitamins (gray), and **FO** full strength MSK medium (yellow). Right shows for comparison nitrite concentration when *N. franklandus* C13 is cultivated in AOA medium over the same time scale (40 days). N=1 per condition.

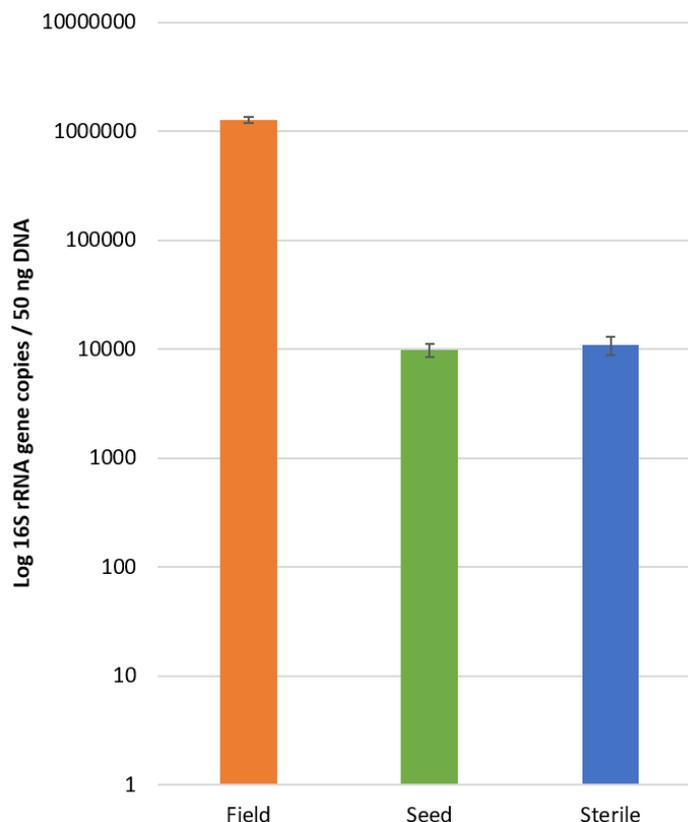


Figure S.3. qPCR experiment performed to compare the absolute abundance of bacterial 16S rRNA gene sequences within root interior for *Triticum aestivum* var. paragon cultivated under sterile conditions to the 16S rRNA gene copy number within previously tested mature field grown plants, and to the wheat seed endosphere. Bars show the mean log 16S rRNA gene copy per 50 ng of DNA extracted from field grown plants (orange), surface sterilised wheat seeds (green), or aseptically cultivated seedlings. N=3 replicate DNA extracts per treatment (for aseptically cultivated plants each extraction was performed on root material pooled from two roots, field and seed extractions were performed as described for those respective experiments). Bars represent  $\pm$  standard error of the mean.

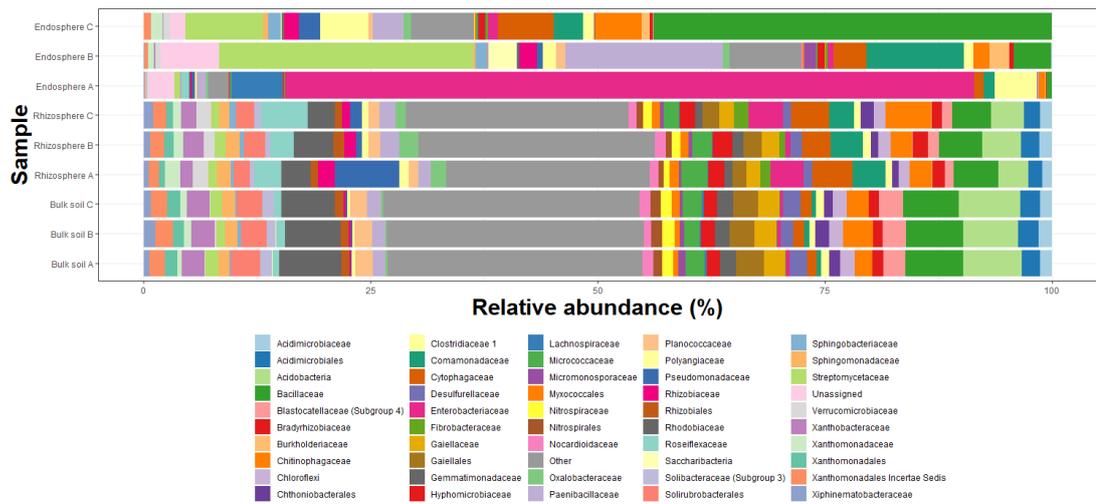


Figure S.2 Metabarcoding performed to profile the unfractionated bacterial communities across three root compartments for plants cultivated in agricultural soil and labelled with  $^{13}\text{CO}_2$  for the stable isotope probing experiment. Bars show the relative abundance (%) of each bacterial taxon within each replicate for the endosphere, rhizosphere or bulk soil of wheat (N=3 replicate plants). Colours indicate different microbial taxa. Within stacked bars taxa are shown in reverse alphabetical order (left to right). The “Other” category includes all taxa with a median relative abundance of 1% or less. ASVs were assigned and are presented to the family level; where family-level taxonomic assignments were unavailable the next highest taxonomic assignment was presented.

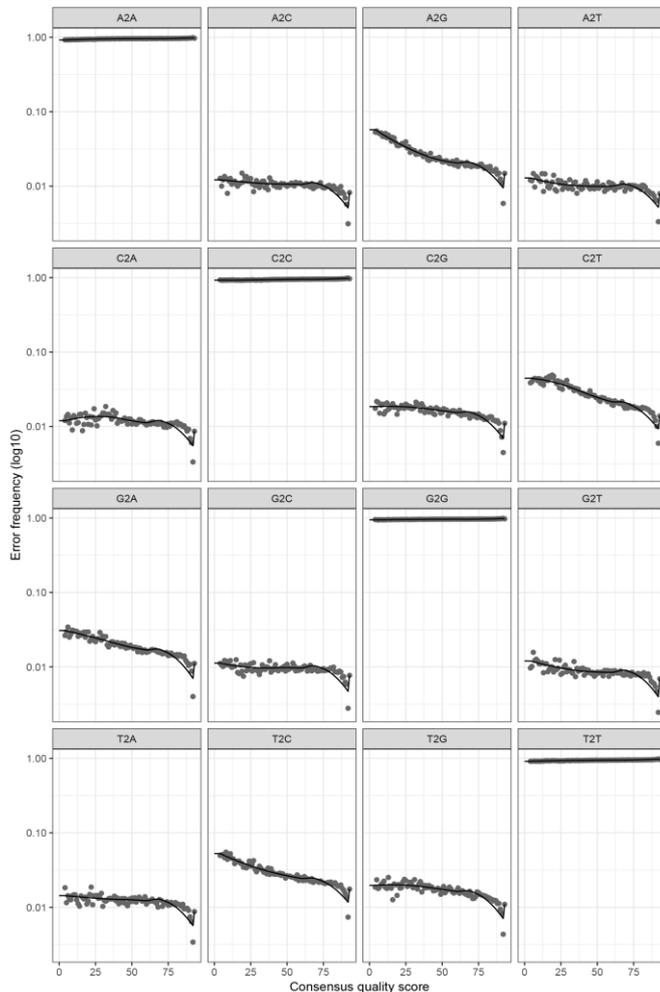


Figure S.5 Quality plot from the PacBio SMRT closed circular sequencing experiment to identify bacterial genera and species within wheat roots. The sample shown is endosphere A, an example of a good quality dataset. Plots show the error frequency (log<sub>10</sub>) over the consensus quality score for each possible nucleotide transition. The black line shows the expected error rate, and the points show the observed error rates. Plots where the observed and expected error rates are similar are indicative of good quality sequencing data.

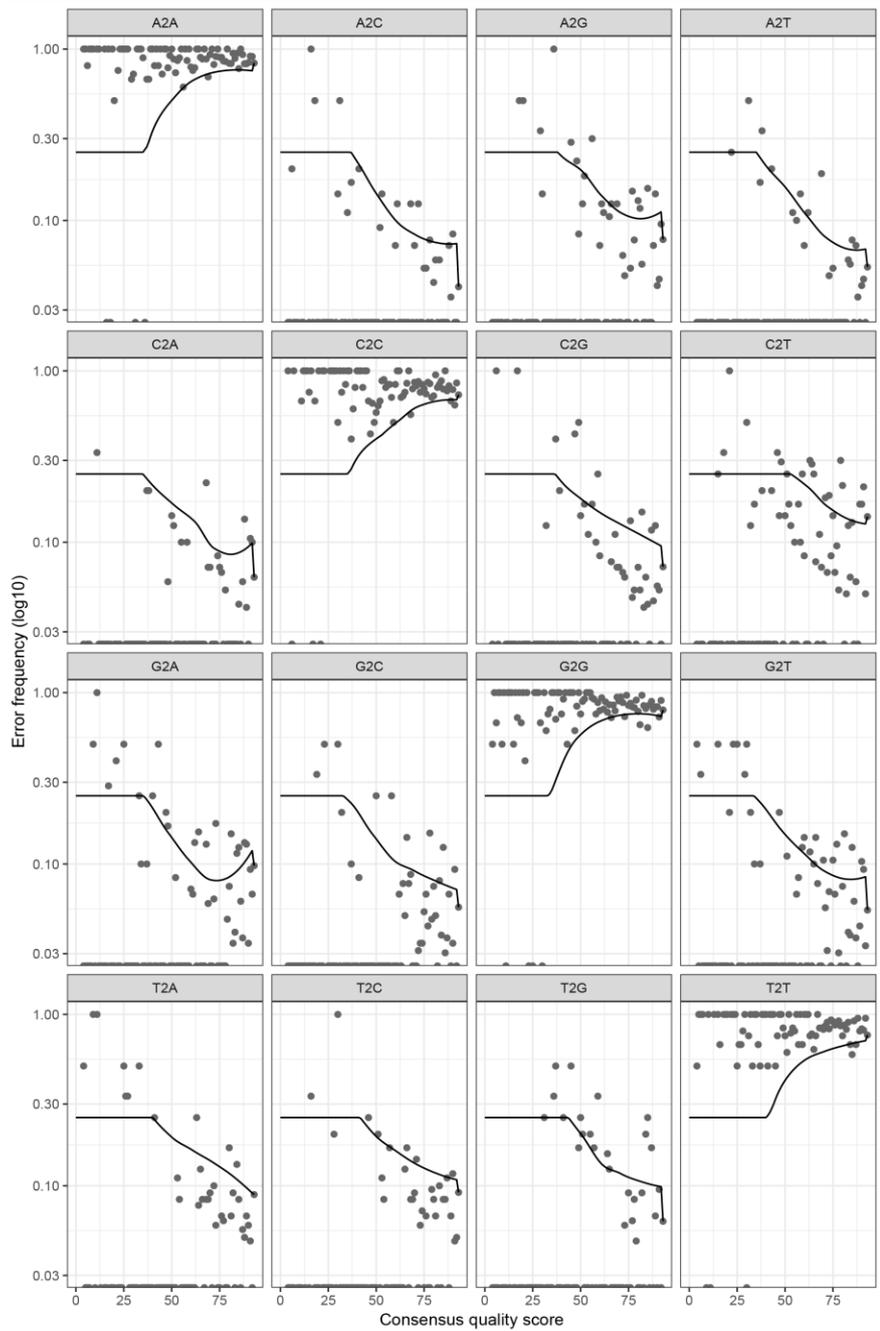


Figure S.6 Quality plot from the PacBio SMRT closed circular sequencing experiment to identify bacterial genera and species within wheat roots. The sample shown is endosphere B, an example of a bad quality dataset. Plots show the error frequency ( $\log_{10}$ ) over the consensus quality score for each possible nucleotide transition. The black line shows the expected error rate, and the points show the observed error rates. Plots where the observed and expected error rates are similar are indicative of good quality sequencing data.

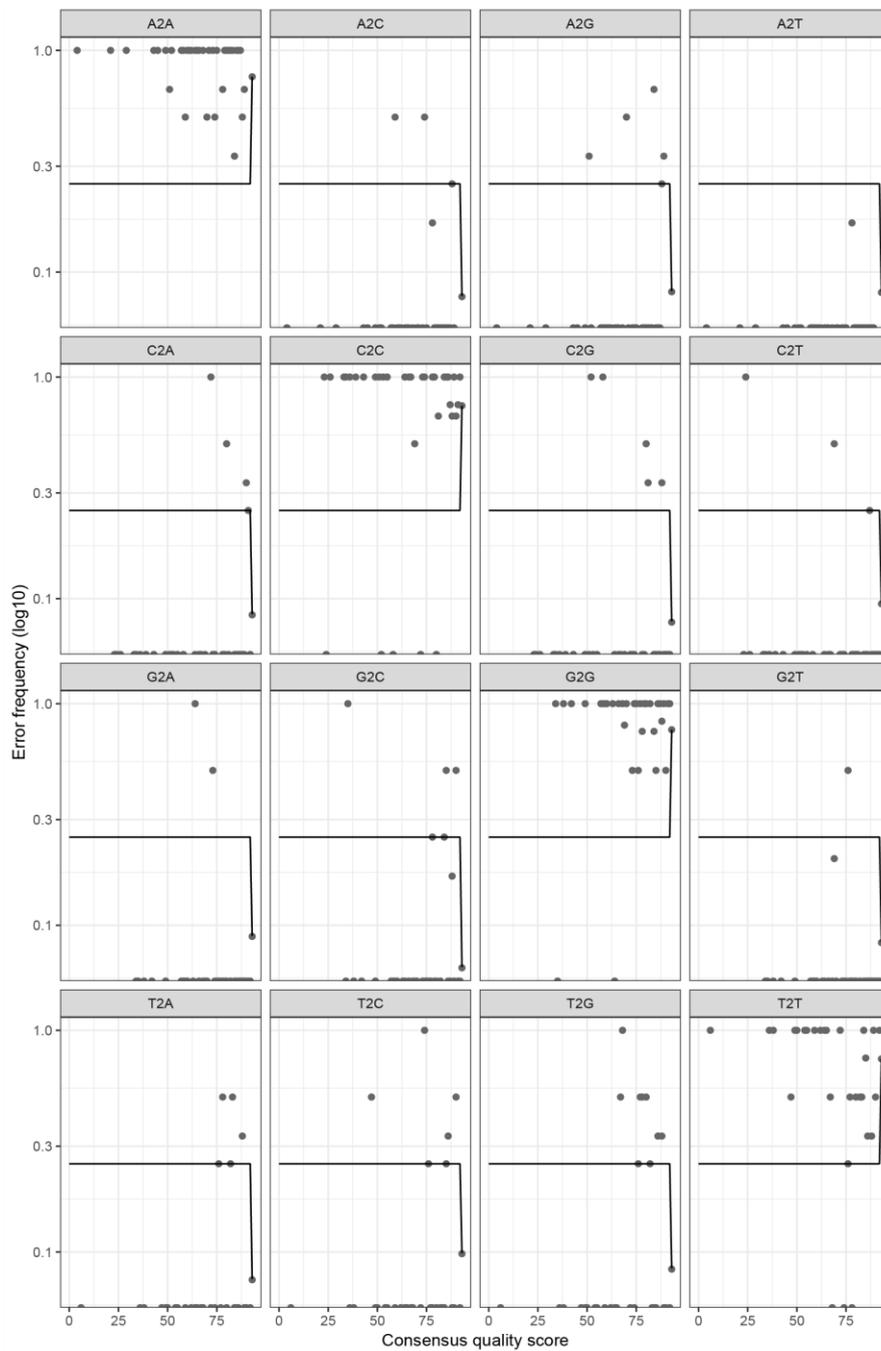


Figure S.7 Quality plot from the PacBio SMRT closed circular sequencing experiment to identify bacterial genera and species within wheat roots. The sample shown is endosphere C, an example of a bad quality dataset. Plots show the error frequency ( $\log_{10}$ ) over the consensus quality score for each possible nucleotide transition. The black line shows the expected error rate, and the points show the observed error rates. Plots where the observed and expected error rates are similar are indicative of good quality sequencing data.

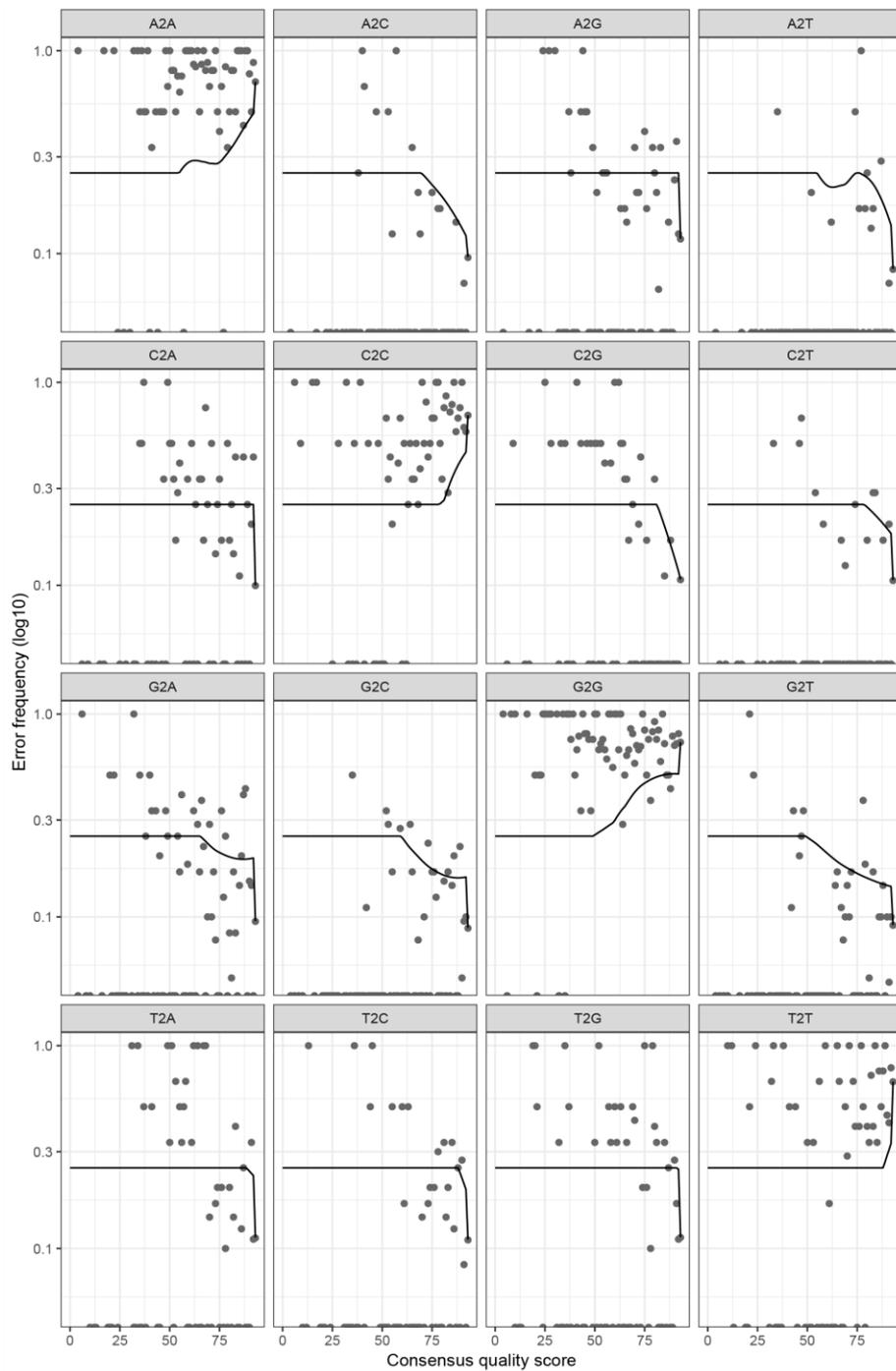


Figure S.8 Quality plot from the PacBio SMRT closed circular sequencing experiment to identify bacterial genera and species within wheat roots. The sample shown is seed A, an example of a bad quality dataset. Plots show the error frequency (log10) over the consensus quality score for each possible nucleotide transition. The black line shows the expected error rate, and the points show the observed error rates. Plots where the observed and expected error rates are similar are indicative of good quality sequencing data.

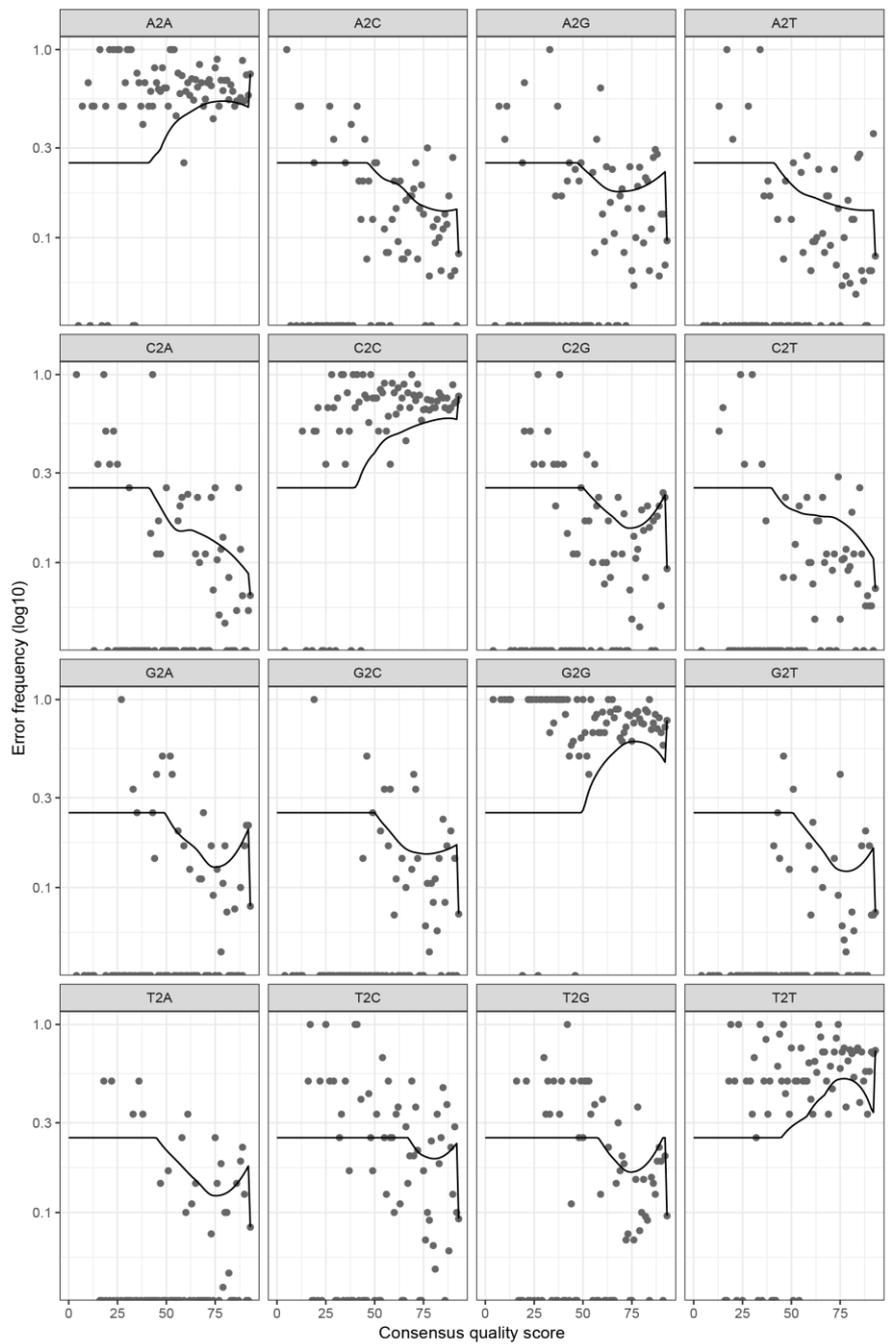


Figure S.9 Quality plot from the PacBio SMRT closed circular sequencing experiment to identify bacterial genera and species within wheat roots. The sample shown is seed B, an example of a bad quality dataset. Plots show the error frequency (log<sub>10</sub>) over the consensus quality score for each possible nucleotide transition. The black line shows the expected error rate, and the points show the observed error rates. Plots where the observed and expected error rates are similar are indicative of good quality sequencing data.

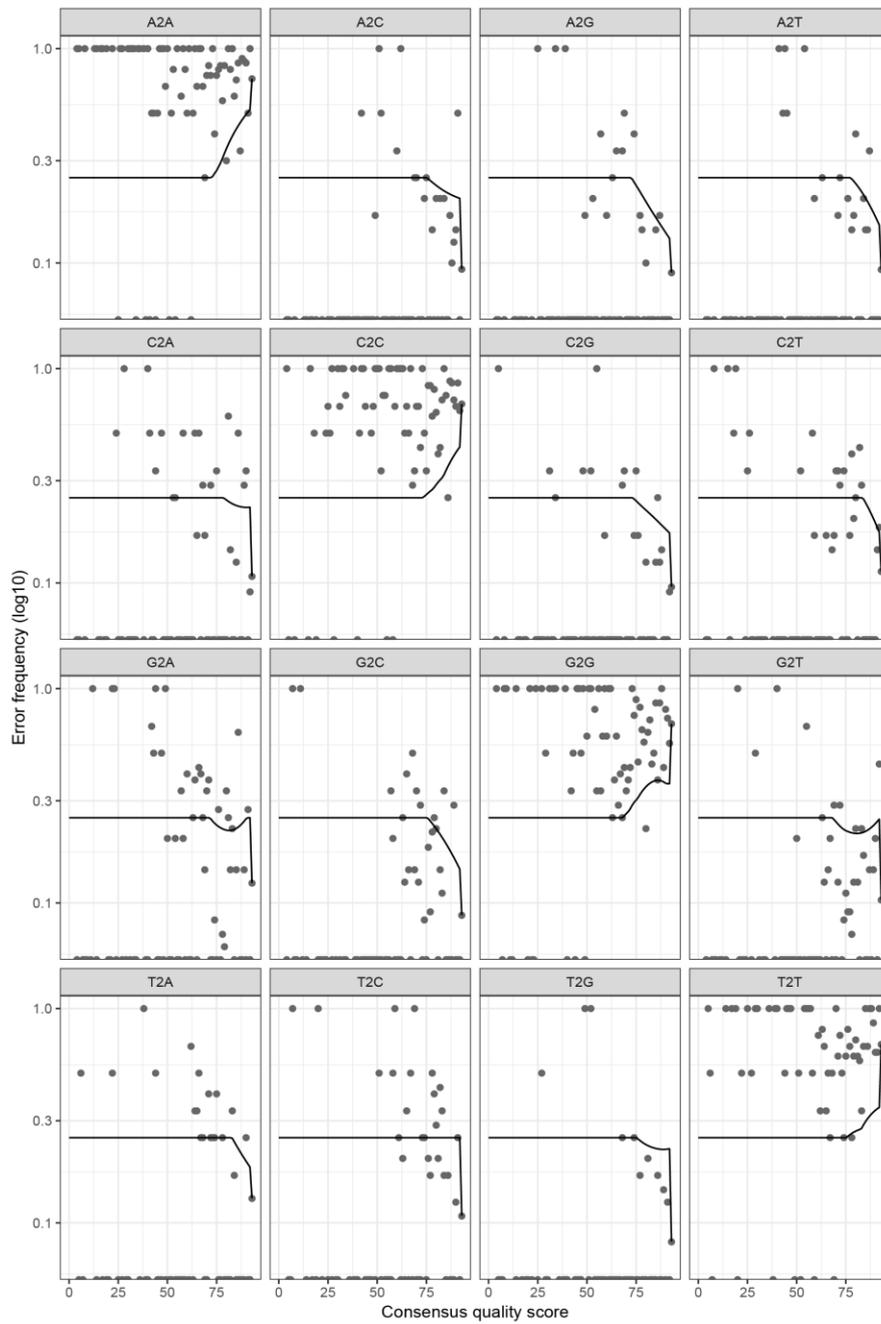


Figure S.10  
 Quality plot from the PacBio SMRT closed circular sequencing experiment to identify bacterial genera and species within wheat roots. The sample shown is seed C, an example of a bad quality dataset. Plots show the error frequency ( $\log_{10}$ ) over the consensus quality score for each possible nucleotide transition. The black line shows the expected error rate, and the points show the observed error rates. Plots where the observed and expected error rates are similar are indicative of good quality sequencing data.

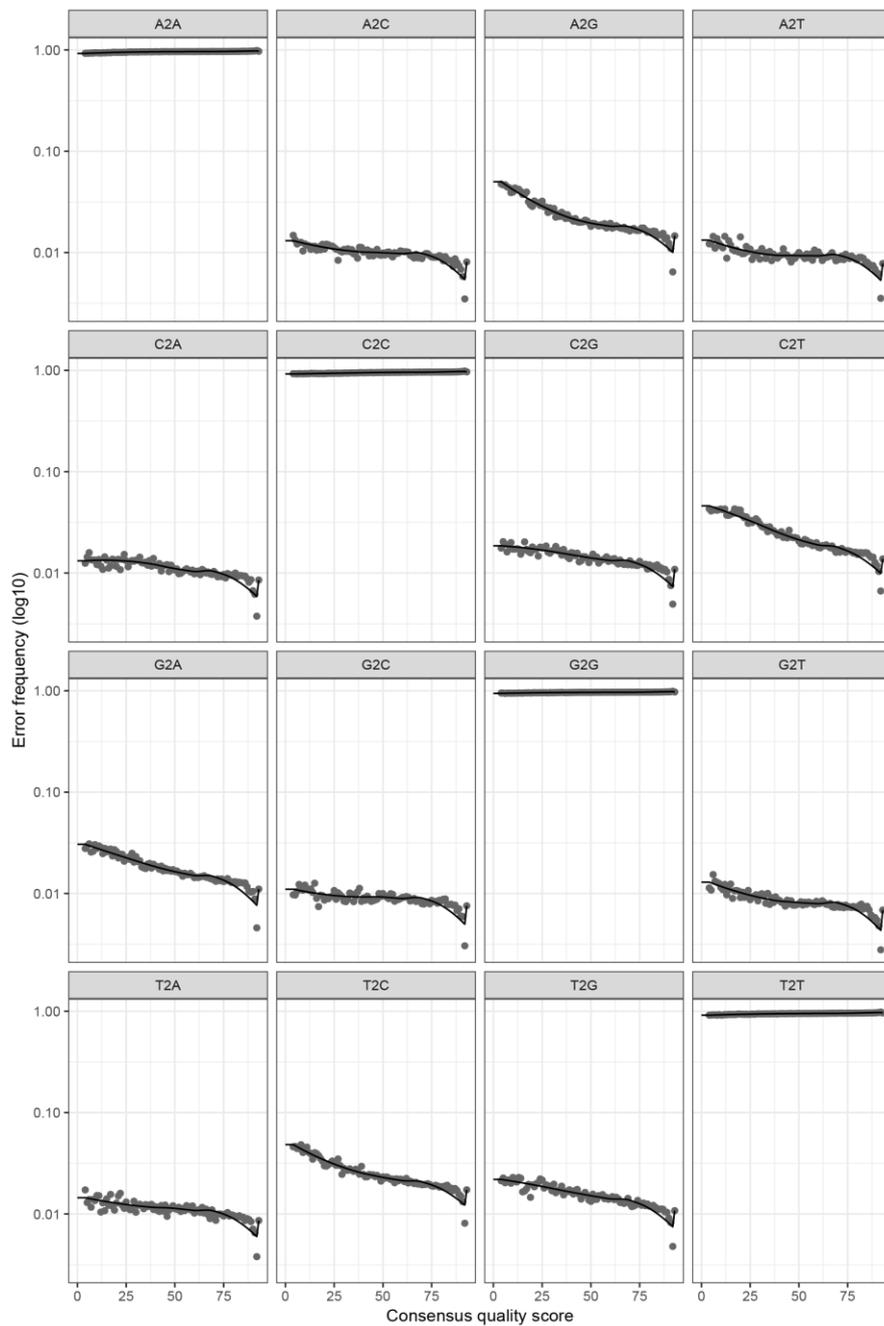


Figure S.11  
 Quality plot from the PacBio SMRT closed circular sequencing experiment to identify bacterial genera and species within wheat roots. Shown is the quality plot for the merged endosphere sample used for analysis. Plots show the error frequency (log10) over the consensus quality score for each possible nucleotide transition. The black line shows the expected error rate, and the points show the observed error rates. Plots where the observed and expected error rates are similar are indicative of good quality

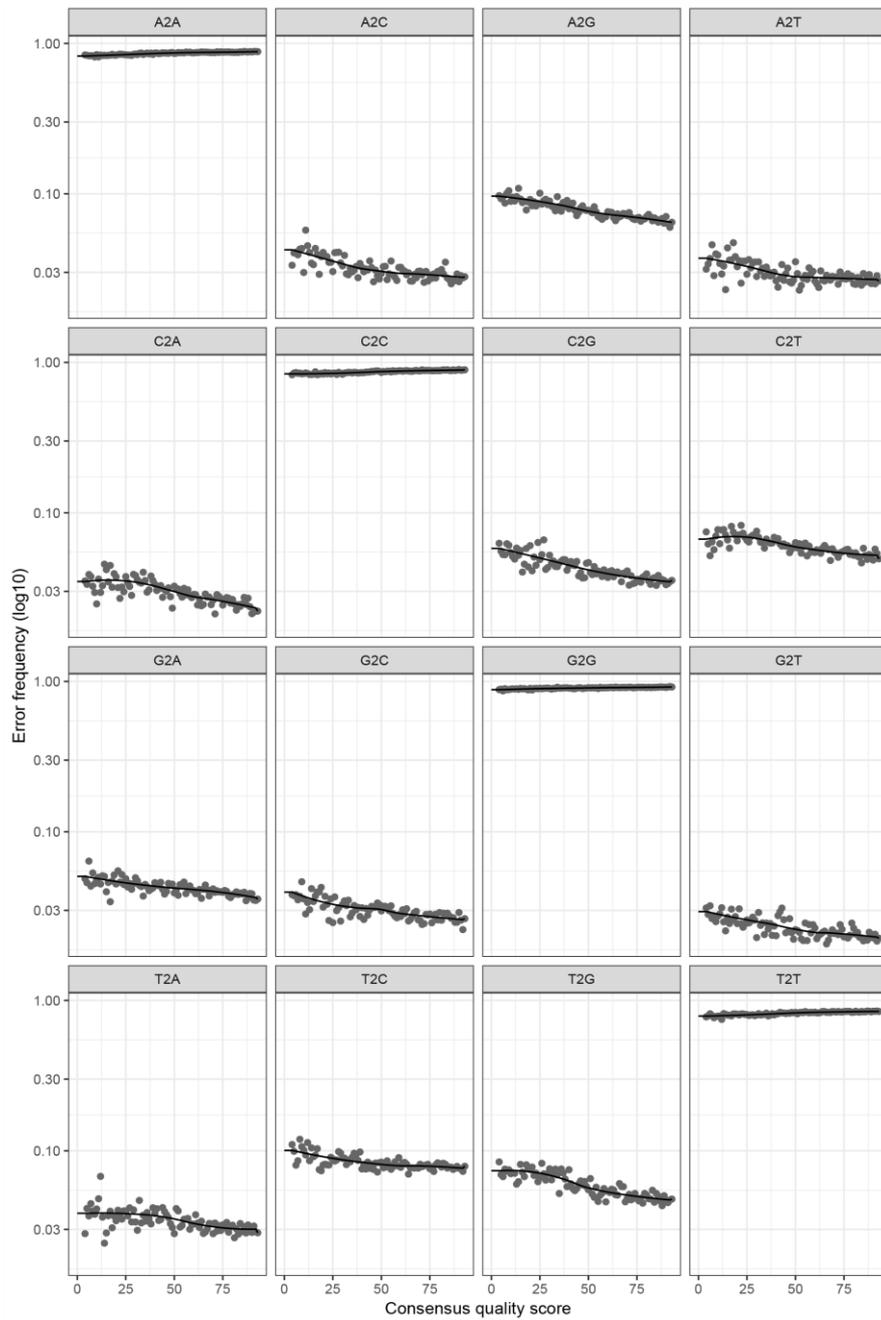


Figure S.12  
 Quality plot from the PacBio SMRT closed circular sequencing experiment to identify bacterial genera and species within wheat roots. Shown is the quality plot for the merged rhizosphere sample used for analysis. Plots show the error frequency (log<sub>10</sub>) over the consensus quality score for each possible nucleotide transition. The black line shows the expected error rate, and the points show the observed error rates. Plots where the observed and expected error rates are similar are indicative of good quality

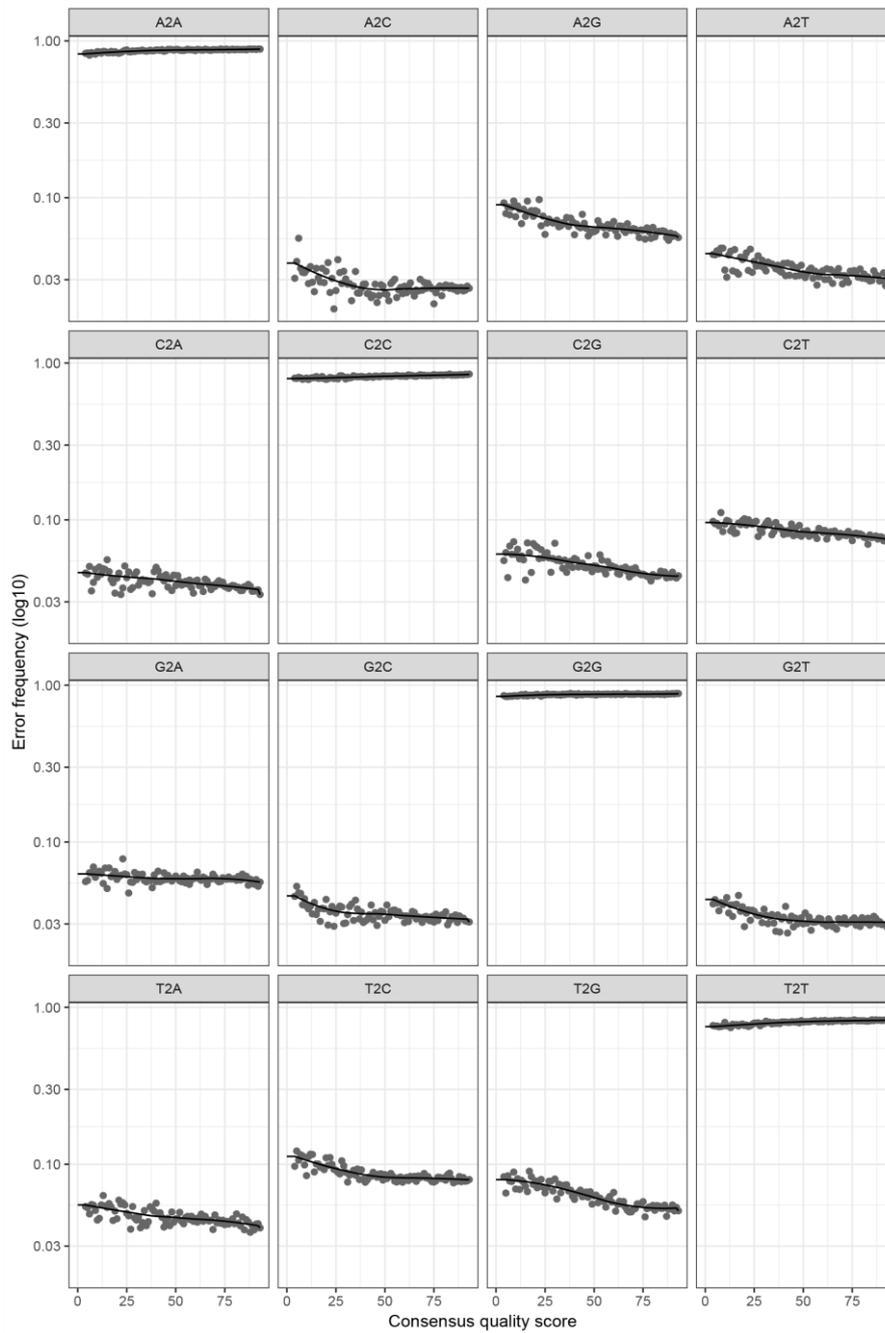


Figure S.12  
 Quality plot from the PacBio SMRT closed circular sequencing experiment to identify bacterial genera and species within wheat roots. Shown is the quality plot for the merged bulk soil samples. Plots show the error frequency (log<sub>10</sub>) over the consensus quality score for each possible nucleotide transition. The black line shows the expected error rate, and the points show the observed error rates. Plots where the observed and expected error rates are similar are indicative of good quality sequencing data.

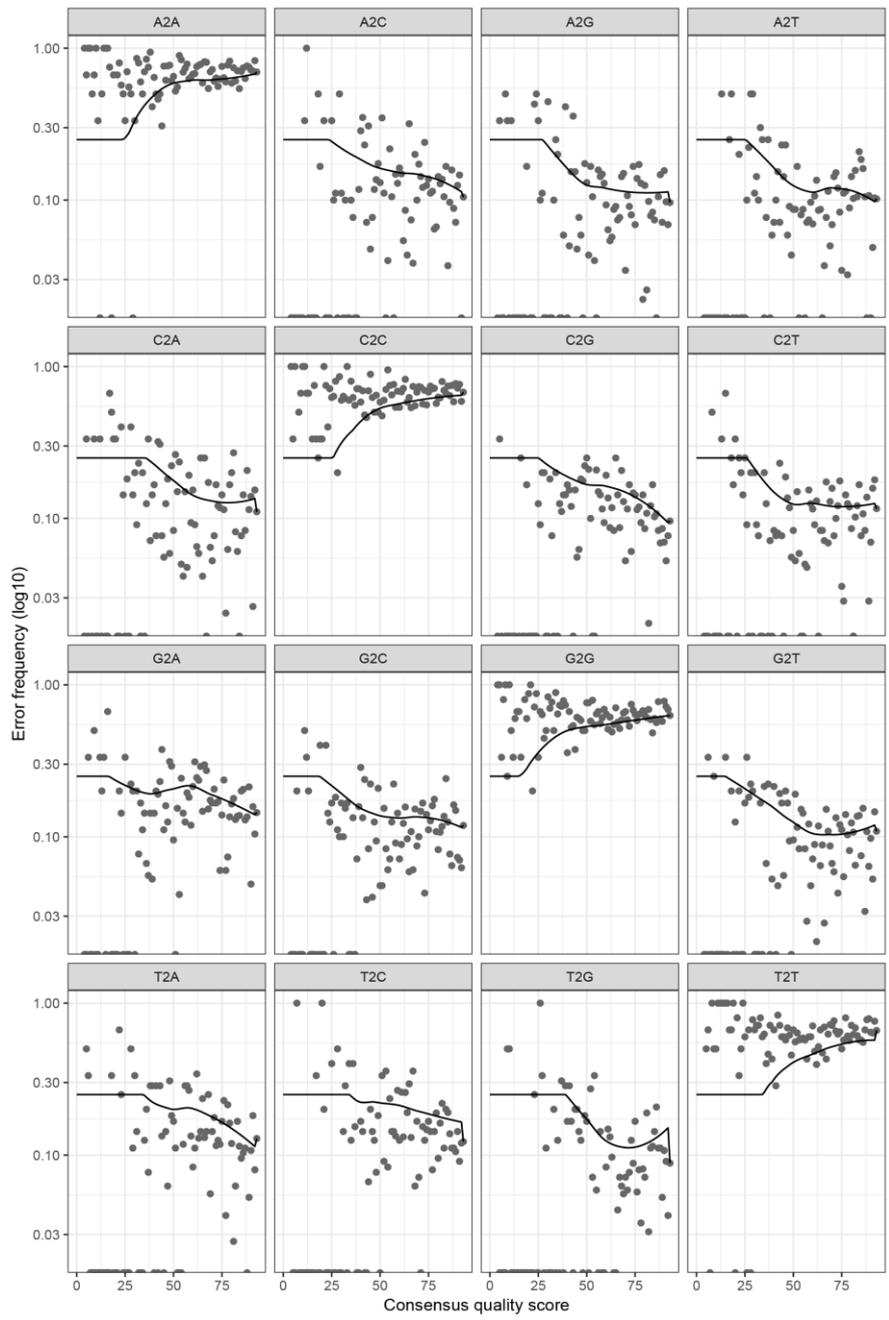


Figure S.13  
 Quality plot from the PacBio SMRT closed circular sequencing experiment to identify bacterial genera and species within wheat roots. Shown is the quality plot for the merged seed samples. Plots show the error frequency (log<sub>10</sub>) over the consensus quality score for each possible nucleotide transition. The black line shows the expected error rate, and the points show the observed error rates. Plots where the observed and expected error rates are similar are indicative of good quality sequencing data.

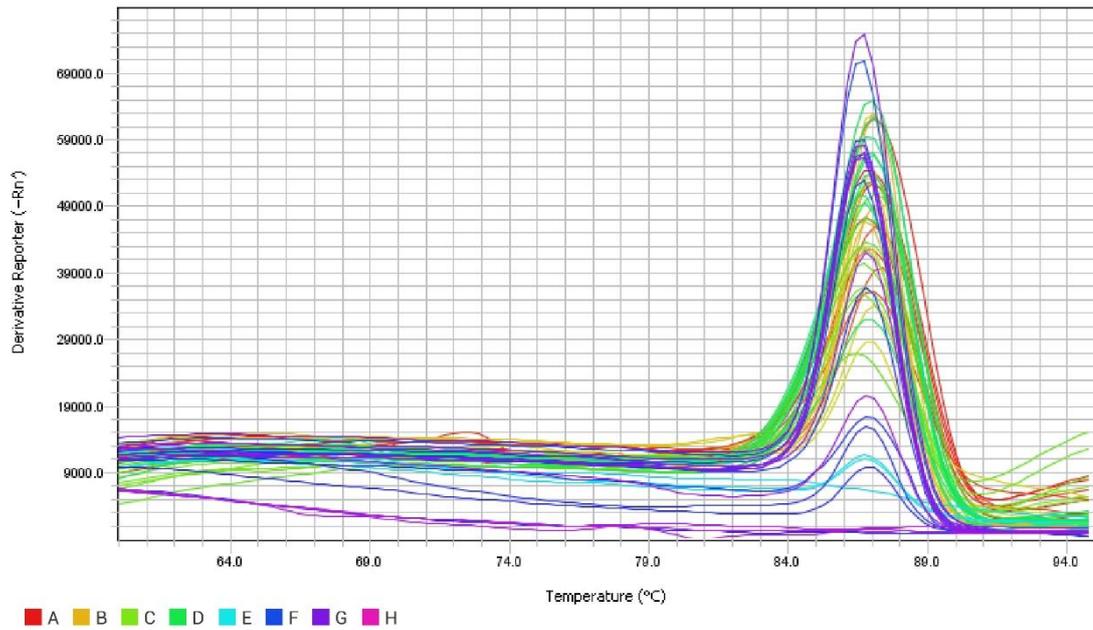


Figure S.14 Melt curve for the archaeal qPCR assay using the primers A771f/A957r to quantify archaeal 16S rRNA gene copy number from soil and root samples, and from fractionated DNA for the stable isotope probing experiment. The peak shows the temperature at which the primers disassociated from the DNA within the samples, expressed as derivative reporter (-RN). Colours show the curve for different sample types. A shows agricultural bulk soil, B shows agricultural rhizosphere, and C shows agricultural endosphere. D shows Levington compost bulk soil, E shows Levington compost rhizosphere, and F shows Levington compost endosphere. G shows the standard curve and H shows a negative control. The graph shows one clear peak at 97-98 degrees Celsius, showing specificity for these primers.

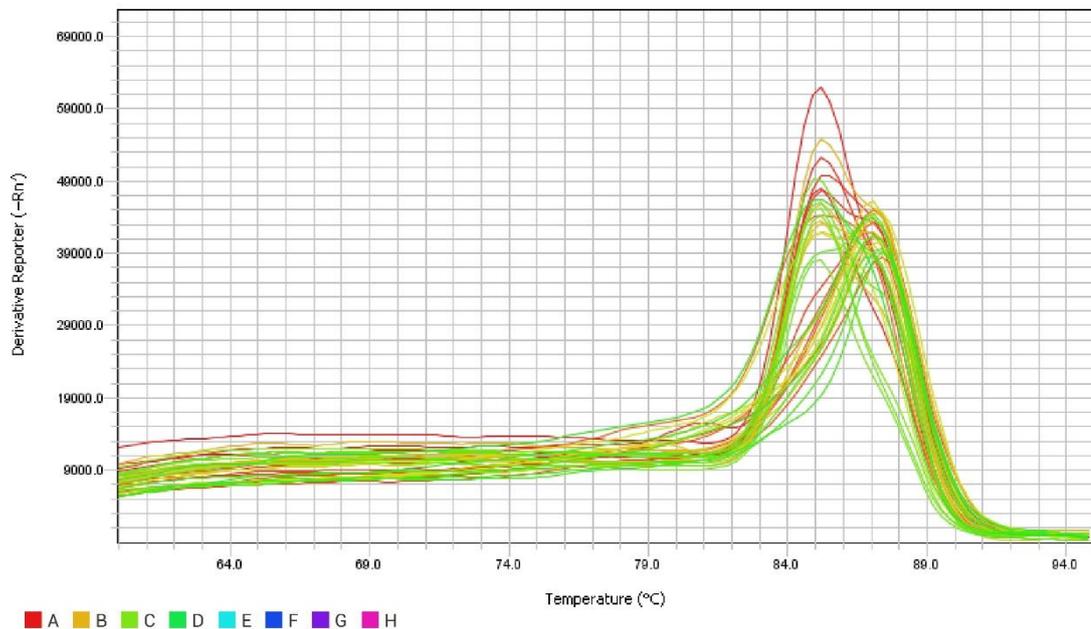


Figure S.15 Melt curve for the bacterial qPCR assay using the primers Com1F/769r to quantify bacterial 16S rRNA gene copy number from soil and root samples. The peak shows the temperature at which the primers disassociated from the DNA within the samples, expressed as derivative reporter (-RN). Colours show the curve for different sample types. A shows agricultural bulk soil, B shows agricultural rhizosphere, and C shows agricultural endosphere. D shows Levington compost bulk soil, E shows Levington compost rhizosphere, and F shows Levington compost endosphere. G shows the standard curve and H shows a negative control. The graph shows a peak at 85-88 degrees Celsius, showing specificity for these primers.

### Melt Curve Plot

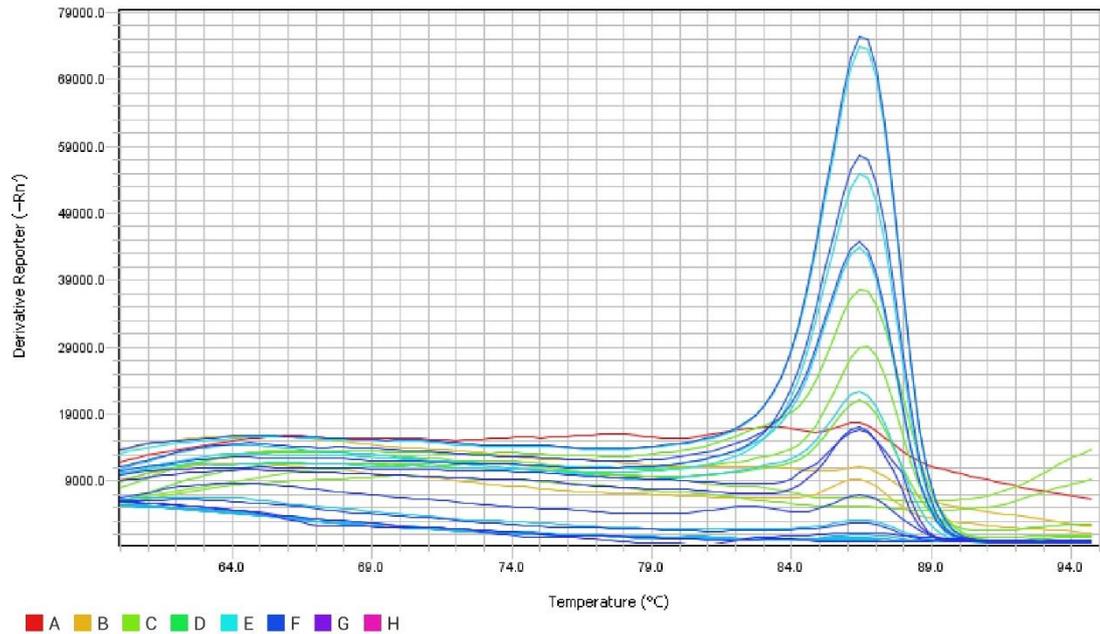


Figure S.16 Melt curve for the fungal qPCR assay using the primers FR1F/FF390R to quantify fungal 18S rRNA gene copy number from soil and root samples, and from fractionated DNA for the stable isotope probing experiment. The peak shows the temperature at which the primers disassociated from the DNA within the samples, expressed as derivative reporter (-RN). Colours show the curve for different sample types. A shows agricultural bulk soil, B shows agricultural rhizosphere, and C shows agricultural endosphere. D shows Levington compost bulk soil, E shows Levington compost rhizosphere, and F shows Levington compost endosphere. G shows the standard curve and H shows a negative control. The graph shows one clear peak at 96 degrees Celsius, showing specificity for these primers.

**Supplementary Table S.1. Qiime2 taxonomy-based filtering stats**

This table shows the total number of trimmed and quality filtered reads recovered for each sample from Illumina amplicon sequencing, and the proportion of reads that were removed by taxonomy based filtering, which was used to remove contaminating chloroplast, mitochondrial and other host-derived sequences.

Experiment	Amplicon	Sample	Total No. trimmed quality filtered reads	No. reads removed by taxonomic filtering	No. reads remaining	% reads discarded
Metabarcoding	A0349F/A0519R Archaeal 16S	BS51.A	68773	475	68298	0.690678
		BS52.A	63866	625	63241	0.978611
		BS53.A	74447	575	73872	0.772362
		BSA1.A	55327	867	54460	1.567047
		BSA2.A	54695	195	54500	0.356523
		BSA3.A	49970	390	49580	0.780468
		BSF1.A	78012	1695	76317	2.172743
		BSF2.A	75560	1892	73668	2.50397
		BSF3.A	78741	1574	77167	1.998959
		BSL1.A	207703	11458	196245	5.516531
		BSL2.A	196347	8785	187562	4.474222
		BSL3.A	133107	4169	128938	3.132067
		E51.A	93810	119	93691	0.126852
		E52.A	92550	711	91839	0.768233
		E53.A	109472	330	109142	0.301447
		EA1.A	79035	54	78981	0.068324
		EA2.A	72672	6124	66548	8.426904
		EA3.A	58574	34	58540	0.058046
		EF1.A	78178	2079	76099	2.659316
		EF2.A	59310	442	58868	0.745237
		EF3.A	57468	856	56612	1.489525
		EL1.A	126246	159	126087	0.125945
		EL2.A	133715	200	133515	0.149572
		RZ51.A	69659	2320	67339	3.33051
		RZ52.A	75587	1549	74038	2.049294
		RZ53.A	68546	1174	67372	1.712718
		RZA1.A	58037	1202	56835	2.071093
		RZA2.A	90071	1176	88895	1.305637
		RZA3.A	68877	1513	67364	2.196669
		RZF1.A	37285	304	36981	0.815341
		RZF2.A	48495	321	48174	0.661924
		RZF3.A	56662	657	56005	1.159507

		RZL1.A	125065	3873	121192	3.09679
		RZL2.A	125751	3886	121865	3.090234
		RZL3.A	123964	2723	121241	2.196605
		EFS1.A	239780	2790	236990	1.1
		EFS2.A	247035	643	246392	0.260287
		EFS3.A	199509	1400	198109	0.701723
		RZFS1.A	247152	881	246271	0.356461
		RZFS2.A	239414	935	238479	0.390537
		RZFS3.A	241137	280	240857	0.116117
		BSFS1.A	254643	226	254417	0.088752
		BSFS2.A	304279	289	303990	0.949786
		BSFS3.A	478179	808	477371	0.168974
	PRK341F/ MPRK806R Bacterial 16S	B.BS51	22970	320	22650	1.393121
		B.BS52	12809	510	12299	3.981575
		B.BS53	22638	396	22242	1.749271
		B.BSA1	13386	455	12931	3.399074
		B.BSA2	18260	778	17482	4.260679
		B.BSA3	17575	977	16598	5.559033
		B.BSL1	32363	3322	29041	10.26481
		B.BSL2	28475	3580	24895	12.57243
		B.BSL3	28042	3478	24564	12.40282
		B.E51	51847	50333	1514	97.07987
		B.E52	46632	46145	487	98.95565
		B.E53	49496	46634	2862	94.21771
		B.EA1	46470	42828	3642	92.16269
		B.EA2	49475	47264	2211	95.53108
		B.EA3	39328	32351	6977	82.25946
		B.EL1	40368	39732	636	98.42449
		B.EL2	46651	46125	526	98.87248
		B.EL3	48154	46843	1311	97.27748
		B.RZ51	17754	976	16778	5.497353
		B.RZ52	18840	983	17857	5.217622
		B.RZ53	17688	782	16906	4.421076
		B.RZA1	22737	687	22050	3.021507
		B.RZA2	23577	813	22764	3.448276
		B.RZA3	20817	647	20170	3.108037
		B.RZL1	21777	632	21145	2.902144
		B.RZL2	26411	1015	25396	3.843096
		B.RZL3	27594	1896	25698	6.871059
		B.BSF1	17020	1101	15919	6.46886
		B.BSF2	15340	568	14772	3.702738
		B.BSF3	19415	414	19001	2.132372

	B.EF1	41256	24671	16585	59.79979
	B.EF2	42388	27888	14500	65.79221
	B.EF3	44801	27308	17493	60.954
	B.RZF1	21359	1201	20158	5.622922
	B.RZF2	24495	472	24023	1.926924
	B.RZF3	19582	760	18822	3.881115
	B.EFS1	53449	138	53311	0.25819
	B.EFS2	60192	350	59842	0.581473
	B.EFS3	54148	755	53393	1.394327
	B.RZFS1	47685	367	47318	0.769341
	B.RZFS2	46770	2122	44648	4.537096
	B.RZFS3	48877	844	48033	1.726784
	B.BSFS1	43114	7244	35870	16.80197
	B.BSFS2	42848	4569	38279	10.66327
	B.BSFS3	48794	6410	42384	13.13686
Varieties experiment	A1E	48221	27979	20242	58.02244
	A2E	121756	74508	47248	61.19452
	A3E	115537	39357	76180	34.06441
PRK341F/ MPRK806R	P1E	101081	33128	67953	32.77372
	P2E	103026	39109	63917	37.96032
	P3E	114165	54378	59787	47.63106
Bacterial 16S	T1E	115064	58114	56950	50.50581
	T2E	105217	43223	61994	41.07986
	T3E	107560	36691	70869	34.11212
A = Axona, P = Paragon, T = Tonic, C = Cadenza, S = Soissons	C1E	107651	50163	57488	46.5978
	C2E	111889	44692	67197	39.94316
	C3E	101231	28929	72302	28.57721
E = Endosphere, R = Rhizosphere, B = Bulk Soil	S1E	91969	50239	41730	54.62602
	S2E	102217	19932	82285	19.49969
	S3E	97174	29967	67207	30.8385
	A1BS	119349	8181	111168	6.854686675
	A1RZ	104196	6434	97762	6.174901148
	A2BS	95272	2738	92534	2.8738769
	A2RZ	123108	4882	118226	3.96562368
	A3BS	117511	14833	102678	12.62264809
	A3RZ	100958	8366	92592	8.286614236
	C1BS	103738	12266	91472	11.8240182
	C1RZ	84120	16441	67679	19.54469805
	C2BS	83249	2217	81032	2.663095052
	C2RZ	108576	3334	105242	3.070660183
	C3BS	90618	2243	88375	2.475225673
	C3RZ	116588	10661	105927	9.14416578

		P1BS	110182	7311	102871	6.63538509
		P1RZ	101856	16821	85035	16.51449105
		P2BS	142507	24251	118256	17.01740967
		P2RZ	125725	6933	118792	5.514416385
		P3BS	122074	16483	105591	13.50246572
		P3RZ	97343	4607	92736	4.732749145
		S1BS	96036	1456	94580	1.51609813
		S1RZ	109993	2094	107899	1.903757512
		S2BS	103877	10073	93804	9.697045544
		S2RZ	106977	2021	104956	1.889191135
		S3BS	95081	3288	91793	3.458104143
		S3RZ	104057	849	103208	0.815898978
		T1BS	97782	6310	91472	6.453130433
		T1RZ	96380	912	95468	0.94625441
		T2BS	93458	3520	89938	3.76639774
		T2RZ	112429	6874	105555	6.114080887
		T3BS	117462	16085	101377	13.69379033
		T3RZ	111907	7950	103957	7.104113237
	fITS7F/ ITS4R	F.BS51	41729	n/a	n/a	n/a
	Fungal ITS2	F.BS52	45565	n/a	n/a	n/a
		F.BS53	46115	n/a	n/a	n/a
		F.BSA1	75779	n/a	n/a	n/a
		F.BSA2	100787	n/a	n/a	n/a
		F.BSA3	94052	n/a	n/a	n/a
		F.BSF1	129866	n/a	n/a	n/a
		F.BSF2	100424	n/a	n/a	n/a
		F.BSF3	81109	n/a	n/a	n/a
		F.BSL1	98179	n/a	n/a	n/a
		F.BSL2	131559	n/a	n/a	n/a
		F.BSL3	120348	n/a	n/a	n/a
		F.E51	101772	n/a	n/a	n/a
		F.E52	114929	n/a	n/a	n/a
		F.E53	44992	n/a	n/a	n/a
		F.EA1	75592	n/a	n/a	n/a
		F.EA2	65668	n/a	n/a	n/a
		F.EA3	89945	n/a	n/a	n/a
		F.EF1	89057	n/a	n/a	n/a
		F.EF2	69637	n/a	n/a	n/a
		F.EF3	102551	n/a	n/a	n/a
		F.RZ51	84654	n/a	n/a	n/a
		F.RZ52	37011	n/a	n/a	n/a
		F.RZ53	31117	n/a	n/a	n/a

		F.RZA1	102095	n/a	n/a	n/a
		F.RZA2	71665	n/a	n/a	n/a
		F.RZA3	84015	n/a	n/a	n/a
		F.RZF1	51408	n/a	n/a	n/a
		F.RZF2	99541	n/a	n/a	n/a
		F.RZF3	45291	n/a	n/a	n/a
		F.RZL1	110595	n/a	n/a	n/a
		F.RZL2	97230	n/a	n/a	n/a
		F.RZL3	81901	n/a	n/a	n/a
		F.EFS1	129650	n/a	n/a	n/a
		F.EFS2	70055	n/a	n/a	n/a
		F.EFS3	76512	n/a	n/a	n/a
		F.RZFS1	56010	n/a	n/a	n/a
		F.RZFS2	64946	n/a	n/a	n/a
		F.RZFS3	36457	n/a	n/a	n/a
		F.BSFS1	104076	n/a	n/a	n/a
		F.BSFS2	71894	n/a	n/a	n/a
		F.BSFS3	80601	n/a	n/a	n/a
Endosphere SIP sequencing (1st run)	PRK341F/ MPRK806R Bacterial 16S	12CHa	93805	43076	50729	45.92079
		12CLa	116684	89134	27550	76.38922
		13CHa	108182	29914	78268	27.65155
		13CLa	97947	48972	48975	49.99847
		12CHb	108089	43142	64947	39.9134
		12CLb	105021	50010	55011	47.61905
		13CHb	95683	73700	21983	77.02518
		13CLb	71120	52215	18905	73.41817
		12CHb	100291	56735	43556	56.57038
		12CLb	92092	84369	7723	91.61382
		13CHb	83840	46040	37800	54.91412
		13CLb	85336	45571	39765	53.40185
Rhizosphere and Bulk soil SIP sequencing		12CHa	13421	885	12536	6.594144
		12CLa	12970	798	12172	6.15266
		13CHa	15960	336	15624	2.105263
		13CLa	12166	142	12024	1.167187
		12CHb	22378	3195	19183	14.27742
		12CLb	14140	986	13154	6.973126
		13CHb	14082	321	13761	2.279506
		13CLb	10732	286	10446	2.664927
		12CHc	15785	316	15469	2.001901
		12CLc	8832	380	8452	4.302536
		13CHc	12383	172	12211	1.389001
		13CLc	10651	146	10505	1.370763

	13CHa	19395	4230	15165	21.80974	
	13CLa	14826	1932	12894	13.03116	
	13CHb	19529	1409	18120	7.214911	
	13CLb	17501	1740	15761	9.942289	
	13CHc	15543	2230	13313	14.34729	
	13CLc	18138	2585	15553	14.25185	
Endosphere SIP sequencing (2nd run)	12CHa	129499	76038	53461	41.28294	
	12CLa	112494	35149	77345	68.75478	
	12CHb	129363	68584	60779	46.9833	
	12CLb	128990	61089	67901	52.64051	
	12CHc	111366	56655	54711	49.1272	
	12CLc	106026	27714	78312	73.86113	
	13CHa	145802	94298	51504	35.32462	
	13CLa	102901	53396	49505	48.10935	
	13CHb	104884	40712	64172	61.18378	
	13CLb	92606	31200	61406	66.30888	
	13CHc	124416	57972	66444	53.40471	
	13CLc	109262	53343	55919	51.17882	
	<p><b>Key:</b> BS = Bulk Soil, RZ = Rhizosphere, E = Endosphere  5 = 50:50 Mix, A = Agricultural Pot, F = Stem Elongation Field, L = Levington F2 Compost, FS = Senescent Field  1/2/3 = sample replicate, A./B./F./ = Archaeal/Bacterial/Fungal amplicon, 12CL – 12C light, 12CH – 12C heavy, 13CL – 13C light, 13CH – 13C heavy, a/b/c – sample replicate.</p>					

The following tables contain the outputs from differential abundance analysis performed using DESeq2. These tables show the significantly differentially abundant taxa for the respective comparisons. Base mean indicates the overall abundance of a family across samples, and log2 fold change (Lfc) the fold change in the abundance of that taxon across comparisons. LfcSE indicates the standard error of that Lfc, and padj shows the adjusted p-value.

**Supplementary Table S.2. DESeq2 outputs, stem elongation field grown**

Differential abundance analysis comparing the abundance of bacterial taxa across compartments for field grown stem elongation growth phase wheat. Comparisons are between abundances within the bulk soil to the rhizosphere, and in the rhizosphere to the endosphere.

<b>Comparison - Bulk Soil-Rhizosphere</b>				
<b>Taxon</b>	<b>Base Mean</b>	<b>Lfc</b>	<b>lfcSE *</b>	<b>padj</b>
<i>Bacillaceae</i>	372.2125	-0.14622	0.239322	0.693852
<i>Bacteriovoraceae</i>	13.98168	2.230435	0.619679	0.002832
<i>Burkholderiaceae</i>	1021.925	0.813465	0.288343	0.029905
<i>Caulobacteraceae</i>	110.498	1.628402	0.494308	0.00759
<i>Chitinophagaceae</i>	165.0146	0.94339	0.293433	0.009317
<i>Devosiaceae</i>	101.8837	2.40246	0.519061	0.000123
<i>Gaiellaceae</i>	258.2943	-0.38662	0.113536	0.006009
<i>Geminococcaceae</i>	26.37856	-0.35258	0.514868	0.640873
<i>Haliangiaceae</i>	26.72952	-0.06375	0.511428	0.948208
<i>Hymenobacteraceae</i>	24.41687	-0.34667	0.477006	0.631582
<i>Ilumatobacteraceae</i>	104.5866	-0.42953	0.354796	0.383108
<i>Microbacteriaceae</i>	280.9652	0.790671	0.218822	0.003023
<i>Micromonosporaceae</i>	70.07192	-0.09121	0.283765	0.849883
<i>Mycobacteriaceae</i>	72.29347	0.013391	0.291328	0.982999
<i>Nitrosomonadaceae</i>	34.0193	-0.54395	0.5127	0.435947
<i>Nitrospiraceae</i>	34.26428	-1.08858	0.331452	0.006069
<i>Opitutaceae</i>	11.38435	1.069774	0.522292	0.094273
<i>Paenibacillaceae</i>	373.3526	0.511014	0.16033	0.009575
<i>Pedosphaeraceae</i>	24.25692	-1.34006	0.45871	0.016595
<i>Planococcaceae</i>	52.96895	-0.19828	0.286047	0.640873
<i>Promicromonosporaceae</i>	141.5928	1.937938	0.499305	0.001299
<i>Pseudomonadaceae</i>	400.3107	1.768233	0.417393	0.000454
<i>Pseudonocardiaceae</i>	34.84908	-0.24604	0.541155	0.773224
<i>Rhizobiaceae</i>	393.09	1.211811	0.182278	2.97E-09
<i>Roseiflexaceae</i>	108.7535	0.144605	0.398015	0.823412

<i>Rubritaleaceae</i>	100.1947	-0.89572	0.244092	0.002699
Saccharimonadaceae	26.26498	1.690537	0.510286	0.00759
Solirubrobacteraceae	129.6696	-0.37406	0.141224	0.044888
Sphingobacteriaceae	776.3987	1.982141	0.452171	0.000292
Spirosomaceae	257.6693	2.757245	0.49784	1.53E-06
Streptomycetaceae	380.0181	0.757702	0.323487	0.076663
Xanthobacteraceae	252.9046	-0.04345	0.295926	0.939636
<b>Comparison – Rhizosphere-Endosphere</b>				
<b>Taxon</b>	<b>Base Mean</b>	<b>Lfc</b>	<b>lfcSE*</b>	<b>padj</b>
Bacillaceae	372.2125	-0.14622	0.239322	0.693852
Bacteriovoracaceae	13.98168	2.230435	0.619679	0.002832
Burkholderiaceae	2152.24	1.142141	0.254425	0.000102
Caulobacteraceae	326.3971	1.281452	0.530015	0.053849
Chitinophagaceae	306.5415	0.988075	0.277398	0.002832
Devosiaceae	310.9739	1.247119	0.265459	4.38E-05
Gaiellaceae	65.93289	-0.57671	0.483931	0.388951
Geminococcaceae	26.37856	-0.35258	0.514868	0.640873
Haliangiaceae	26.72952	-0.06375	0.511428	0.948208
Hymenobacteraceae	24.41687	-0.34667	0.477006	0.631582
Ilumatobacteraceae	104.5866	-0.42953	0.354796	0.383108
Microbacteriaceae	364.1172	0.699782	0.316859	0.068361
Micromonosporaceae	243.0461	1.932803	0.296194	3.39E-09
Mycobacteriaceae	72.29347	0.013391	0.291328	0.982999
Nitrosomonadaceae	34.0193	-0.54395	0.5127	0.435947
Nitrospiraceae	34.26428	-1.08858	0.331452	0.006069
Opitutaceae	11.38435	1.069774	0.522292	0.094273
Paenibacillaceae	367.5622	0.532595	0.241368	0.068361
Pedosphaeraceae	24.25692	-1.34006	0.45871	0.016595
Planococcaceae	52.96895	-0.19828	0.286047	0.640873
Promicromonosporaceae	631.1794	1.583836	0.370707	0.000242
Pseudomonadaceae	337.4026	-0.05043	0.471101	0.952873
Pseudonocardiaceae	314.0513	2.733577	0.434434	7.82E-09
Rhizobiaceae	503.9776	0.568904	0.274098	0.090323
Roseiflexaceae	108.7535	0.144605	0.398015	0.823412
Rubritaleaceae	140.9361	0.749244	0.331561	0.062729
Saccharimonadaceae	115.8647	1.595737	0.55365	0.017169
Solirubrobacteraceae	37.03626	-0.3467	0.346116	0.465436
Sphingobacteriaceae	1388.722	0.797741	0.338137	0.054036
Spirosomaceae	607.5499	1.017229	0.414429	0.052248
Streptomycetaceae	4613.887	2.536194	0.400799	7.82E-09
Xanthobacteraceae	252.9046	-0.04345	0.295926	0.939636
*lfcSE= log2 Fold Change Standard Error Lfc = log2 Fold Change				

**Supplementary Table S.3. DESeq2 outputs (all pot grown and stem elongation)-**

Differential abundance analysis comparing the abundance of bacterial taxa across compartments across all conditions. Comparisons use the pooled data for pot cultivated plants (Levington compost, agricultural soil, and the 50:50 mix) and field cultivated stem elongation growth phase plants. Comparisons were between abundances within the bulk soil to the rhizosphere, and in the rhizosphere to the endosphere.

<b>Comparison - Bulk Soil-Rhizosphere</b>				
<b>Taxon</b>	<b>Base Mean</b>	<b>Lfc</b>	<b>lfcSE *</b>	<b>padj</b>
Unknowns	2068.626	-1.19715	0.411885	0.026106
<i>Burkholderiaceae</i>	985.327	3.179403	0.485224	2.83E-09
Uncultured	756.9371	-1.25519	0.465714	0.046897
Unassigned	638.8419	-1.13566	0.296593	0.001838
<i>Rhizobiaceae</i>	308.3923	1.995953	0.319026	9.85E-09
<i>Streptomycetaceae</i>	292.7344	1.160709	0.344891	0.008221
<i>Pseudomonadaceae</i>	267.3359	4.65929	0.654589	1.10E-10
<i>Rubritaleaceae</i>	184.9994	3.54871	0.808995	0.00023
env.OPS 17	136.8828	-2.13424	0.673093	0.012669
<i>Haliangiaceae</i>	80.43715	-0.94404	0.278696	0.008221
<i>Spirosomaceae</i>	59.57213	4.777798	0.742664	4.16E-09
<i>Pyrinomonadaceae</i>	58.62283	-2.1681	0.648117	0.008221
<i>Fibrobacteraceae</i>	57.28673	2.684165	0.854304	0.012911
metagenome	54.20017	-1.47081	0.382967	0.001838
<i>Cellvibrionaceae</i>	25.5762	2.138713	0.668686	0.012564
<b>Comparison – Rhizosphere-Endosphere</b>				
<b>Taxon</b>	<b>Base Mean</b>	<b>Lfc</b>	<b>lfcSE*</b>	<b>padj</b>
<i>Streptomycetaceae</i>	1540.243	-4.577	0.532667	7.14E-16
<i>Erysipelotrichaceae</i>	7.280075	-6.11435	0.737884	4.91E-15
<i>Thermomonosporaceae</i>	14.85748	-4.76837	0.696907	2.18E-10
<i>Pseudonocardiaceae</i>	88.20983	-4.75912	0.809506	8.67E-08
<i>Pedosphaeraceae</i>	14.18555	2.540781	0.498724	5.87E-06
<i>Polyangiaceae</i>	161.9785	-3.4803	0.769874	8.63E-05
<i>Chitinophagaceae</i>	424.247	-3.02177	0.705147	0.000192
<i>Rhizobiaceae</i>	267.3061	-1.11239	0.258018	0.000192
<i>Promicromonosporaceae</i>	88.51291	-3.44988	0.844883	0.000414
Unassigned	78.08997	1.719875	0.424449	0.000427
<i>Gemmatimonadaceae</i>	70.10253	2.198493	0.561945	0.000698
<i>Solirubrobacteraceae</i>	19.04678	1.762621	0.490947	0.002313
<i>Paenibacillaceae</i>	123.1825	-2.2608	0.650672	0.003306
WD2101 soil group	30.91117	2.731383	0.849034	0.007771
<i>Cellvibrionaceae</i>	27.81366	-2.00047	0.628252	0.008129
<i>Cytophagaceae</i>	14.87105	-1.99234	0.644695	0.010495

<i>Microscillaceae</i>	152.3535	-2.09948	0.719648	0.017442
uncultured actinobacterium	8.533708	1.928468	0.684966	0.022733
<i>Saccharimonadaceae</i>	21.08344	-2.26986	0.819863	0.02489
CPla-3 termite group	11.36798	2.312846	0.905758	0.044792
*lfsSE= log2 Fold Change Standard Error, Lfc = log2 Fold Change				

#### Supplementary Table S.4. DESeq2 outputs senescent plants

Differential abundance analysis comparing the abundance of bacterial taxa across compartments for field grown wheat after senescence. Comparisons are between abundances within the bulk soil to the rhizosphere, and in the rhizosphere to the endosphere.

Comparison - Bulk Soil-Rhizosphere				
Taxon				
	Base Mean	Lfc	lfcSE*	padj
<i>Bacillaceae</i>	767.3759	0.075595	0.050122	0.26299
<i>Bacteriovoraceae</i>	355.3372	0.446428	0.06923	1.13E-08
<i>Burkholderiaceae</i>	1720.065	0.178619	0.035847	1.25E-05
<i>Chitinophagaceae</i>	161.8836	-0.19486	0.11932	0.227669
<i>Devosiaceae</i>	2062.64	0.028955	0.034216	0.521632
<i>Gaiellaceae</i>	91.78086	0.204319	0.204316	0.466624
<i>Geminococcaceae</i>	390.196	0.238407	0.092834	0.032987
<i>Gemmataceae</i>	283.8183	0.288585	0.099276	0.01521
<i>Haliangiaceae</i>	304.0407	0.205954	0.083663	0.043212
<i>Ilumatobacteraceae</i>	430.3705	0.249691	0.071967	0.003725
<i>Microbacteriaceae</i>	576.283	0.137523	0.072957	0.156406
<i>Micromonosporaceae</i>	235.302	0.068233	0.097531	0.576395
<i>Mycobacteriaceae</i>	237.3228	0.059733	0.090561	0.592464
<i>Nitrosomonadaceae</i>	211.5452	-0.04304	0.122052	0.775018
<i>Nitrospiraceae</i>	631.786	0.202049	0.060474	0.005216
<i>Opitutaceae</i>	348.0111	0.019542	0.069559	0.812289
<i>Paenibacillaceae</i>	516.4374	-0.19252	0.062091	0.010162
<i>Pedosphaeraceae</i>	486.8574	-0.28646	0.072825	0.000837
<i>Planococcaceae</i>	445.4718	0.147492	0.083526	0.184347
<i>Pseudomonadaceae</i>	347.0514	-0.26987	0.101567	0.027178
<i>Pseudonocardiaceae</i>	166.4812	0.286513	0.147584	0.142084
<i>Rhizobiaceae</i>	103.1054	-0.48027	0.186006	0.032744
<i>Rhodomicrobiaceae</i>	540.1628	0.190274	0.061999	0.010739
<i>Roseiflexaceae</i>	222.2028	0.566638	0.102377	1.15E-06
<i>Rubritaleaceae</i>	246.7347	0.436395	0.133666	0.006443
<i>Saccharimonadaceae</i>	404.6714	0.07716	0.106963	0.567085
<i>Solirubrobacteraceae</i>	214.7912	0.164709	0.105573	0.247343
<i>Sphingobacteriaceae</i>	153.0475	0.174746	0.163668	0.432822
<i>Spirosomaceae</i>	319.3659	-0.48343	0.119487	0.000586

<i>Streptomycetaceae</i>	163.2731	-0.09857	0.12104	0.521632
<i>Xanthobacteraceae</i>	189.924	-0.12651	0.11795	0.432822
<b>Comparison – Rhizosphere-Endosphere</b>	360.5323	-0.25971	0.072029	0.002396
<b>Taxon</b>				
<i>Bacillaceae</i>	<b>Base Mean</b>	<b>Lfc</b>	<b>lfcSE*</b>	<b>padj</b>
<i>Bacteriovoracaceae</i>	669.8586	-0.24338	0.093574	0.02817
<i>Burkholderiaceae</i>	211.6517	-0.38527	0.16022	0.0426
<i>Caulobacteraceae</i>	1560.524	-0.05437	0.071256	0.556862
<i>Chitinophagaceae</i>	161.1426	-0.32272	0.179428	0.126463
<i>Gaiellaceae</i>	2810.665	0.334696	0.09971	0.003625
<i>Geminococcaceae</i>	616.729	0.668651	0.14625	5.37E-05
<i>Gemmataceae</i>	253.4938	0.051528	0.157662	0.835729
<i>Haliangiaceae</i>	609.9269	0.873427	0.10693	1.56E-14
<i>Hymenobacteraceae</i>	412.6111	0.106914	0.097998	0.393264
<i>Ilumatobacteraceae</i>	421.6195	0.664185	0.119973	6.19E-07
<i>Microbacteriaceae</i>	1702.794	1.175225	0.140792	6.99E-15
<i>Micromonosporaceae</i>	174.5953	-0.56967	0.190141	0.01013
<i>Mycobacteriaceae</i>	251.2053	0.040195	0.145248	0.839919
<i>Nitrosomonadaceae</i>	430.8715	0.710153	0.17698	0.000429
<i>Nitrospiraceae</i>	687.0426	0.216746	0.141276	0.19838
<i>Opitutaceae</i>	613.7027	0.610044	0.137464	8.26E-05
<i>Paenibacillaceae</i>	1042.123	0.607463	0.131676	4.95E-05
<i>Pedosphaeraceae</i>	959.6071	0.525368	0.079788	1.52E-09
<i>Planococcaceae</i>	677.0895	0.550655	0.095806	2.26E-07
<i>Pseudomonadaceae</i>	369.4337	-0.28308	0.191611	0.214734
<i>Pseudonocardiaceae</i>	195.1522	0.386658	0.234509	0.165314
<i>Rhizobiaceae</i>	222.4399	0.516128	0.159337	0.005211
<i>Rhodomicrobiaceae</i>	385.9288	-0.45277	0.114062	0.00048
<i>Roseiflexaceae</i>	100.4318	-0.77128	0.17287	8.13E-05
<i>Rubritaleaceae</i>	144.6883	-0.43948	0.214584	0.08816
<i>Saccharimonadaceae</i>	366.9267	-0.17432	0.175362	0.444729
<i>Solirubrobacteraceae</i>	209.715	0.035154	0.126667	0.839919
<i>Sphingobacteriaceae</i>	140.2452	-0.04875	0.182615	0.839919
<i>Spirosomaceae</i>	599.1594	0.360183	0.106568	0.003625
<i>Streptomycetaceae</i>	187.2959	-0.00022	0.28912	0.999402
<i>Xanthobacteraceae</i>	371.4646	0.613746	0.142608	0.00014
*lfcSE= log2 Fold Change Standard Error, Lfc = log2 Fold Change	370.2052	-0.33994	0.144082	0.04694

**Supplementary Table S.5. DESeq2 outputs stem elongation/senescent plants**

Differential abundance analysis comparing the abundance of bacterial taxa within the endosphere of senesced plants compared to stem elongation growth phase plants.

<b>Comparison – Senescent endosphere/Stem elongation endosphere</b>				
<b>Taxon</b>	<b>Base Mean</b>	<b>Lfc</b>	<b>lfcSE*</b>	<b>padj</b>
<i>Bacillaceae</i>	357.5098	-0.81595	0.241208	0.004485
<i>Bacteriovoraceae</i>	25.6759	-1.66011	0.594842	0.019471
<i>Burkholderiaceae</i>	3318.41	-1.25399	0.238716	1.66E-06
<i>Caulobacteraceae</i>	541.1063	-0.79572	0.360076	0.072479
<i>Chitinophagaceae</i>	870.3272	0.176643	0.254357	0.587213
<i>Devosiaceae</i>	528.9238	-1.11192	0.28085	0.000579
<i>Gaiellaceae</i>	68.69234	0.214398	0.45143	0.69004
<i>Geminococcaceae</i>	38.29272	0.370222	0.423023	0.515508
<i>Haliangiaceae</i>	47.10503	0.248007	0.470804	0.657529
<i>Hymenobacteraceae</i>	19.25489	-0.84456	0.475021	0.157109
<i>Ilumatobacteraceae</i>	160.0597	0.364665	0.301336	0.395293
<i>Microbacteriaceae</i>	656.3924	-0.31577	0.245576	0.354469
<i>Micromonosporaceae</i>	1086.754	-1.17346	0.357779	0.005467
<i>Mycobacteriaceae</i>	84.1459	-0.47063	0.259804	0.149081
<i>Nitrosomonadaceae</i>	43.46863	0.355034	0.473786	0.567054
<i>Nitrospiraceae</i>	33.82162	0.659862	0.322814	0.090987
<i>Opitutaceae</i>	22.22379	-0.56716	0.489488	0.407432
<i>Paenibacillaceae</i>	431.831	-1.74822	0.248606	5.09E-11
<i>Pedosphaeraceae</i>	18.52983	0.736093	0.466216	0.219935
<i>Planococcaceae</i>	59.45307	-0.31161	0.280003	0.407432
<i>Promicromonosporaceae</i>	1086.754	-1.17346	0.357779	0.005467
<i>Pseudomonadaceae</i>	304.2958	-0.94093	0.341331	0.020137
<i>Pseudonocardiaceae</i>	878.523	-0.10126	0.262648	0.73668
<i>Rhizobiaceae</i>	1302.42	0.255903	0.231463	0.407432
<i>Rhodomicrobiaceae</i>	5.325793	0.713473	0.635217	0.407432
<i>Roseiflexaceae</i>	104.1034	-1.21258	0.403322	0.010405
<i>Rubritaleaceae</i>	238.5271	-0.68881	0.318014	0.075784
<i>Saccharimonadaceae</i>	613.9465	0.60443	0.56084	0.41666
<i>Solirubrobacteraceae</i>	40.43118	-0.13417	0.411122	0.775156
<i>Sphingobacteriaceae</i>	1977.09	-1.36091	0.330803	0.000324
<i>Spirosomaceae</i>	815.2618	-1.9214	0.301822	3.23E-09
<i>Streptomycetaceae</i>	7007.769	-2.15252	0.314852	1.62E-10
<i>Xanthobacteraceae</i>	424.1464	0.19374	0.240251	0.554194

\*lfcSE= log2 Fold Change Standard Error, Lfc = log2 Fold Change

**Supplementary Table S.6. DESeq2 outputs Significantly differentially abundant fungal taxa**

Differential abundance analysis comparing the abundance of fungal taxa across compartments for either senescent or stem elongation growth phase plants. Comparisons are between the abundances in the bulk soil to the rhizosphere, and the rhizosphere to the endosphere.

<b>Comparison - Bulk Soil-Rhizosphere Senescent</b>				
<b>Taxon</b>	<b>Base Mean</b>	<b>Lfc</b>	<b>lfcSE*</b>	<b>padj</b>
<i>Ambisporaceae</i>	136.0381	-7.31509	1.504427	3.60E-05
<i>Chaetosphaeriaceae</i>	583.7226	-4.90418	1.272226	0.000891
<i>Cladosporiaceae</i>	1751.722	0.352167	0.865945	0.861533
<i>Erythrobasidiaceae</i>	88.04139	-0.04082	0.908122	0.964151
<i>Hypocreales Incertae sedis</i>	11149.53	-0.45642	0.716685	0.774593
<i>Mortierellaceae</i>	3301.221	-2.72388	0.842895	0.004771
<i>Sporidiobolaceae</i>	535.6329	3.439894	0.856549	0.000612
<i>Sydowiellaceae</i>	524.9519	-0.50637	0.762741	0.774593
<i>Parmeliaceae</i>	40.84577	5.768819	1.221575	3.61E-05
<b>Comparison – Rhizosphere-Endosphere Senescent</b>				
<b>Taxon</b>	<b>Base Mean</b>	<b>Lfc</b>	<b>lfcSE*</b>	<b>padj</b>
<i>Ambisporaceae</i>	1.053667	1.051434	2.149638	0.756691
<i>Chaetosphaeriaceae</i>	43.17347	2.16618	0.73166	0.019034
<i>Cladosporiaceae</i>	438.7817	-1.79186	0.722699	0.049133
<i>Erythrobasidiaceae</i>	19.80491	-4.76126	1.234989	0.001089
<i>Hypocreales Incertae sedis</i>	2296.368	-2.83945	0.668941	0.000339
<i>Mortierellaceae</i>	225.3781	-1.99924	0.697942	0.021581
<i>Parmeliaceae</i>	15.73176	1.890016	1.11563	0.233124
<i>Sporidiobolaceae</i>	191.0621	-3.80797	0.717967	3.52E-06
<i>Sydowiellaceae</i>	117.4911	-1.88265	0.696616	0.030472
<b>Comparison – Bulk Soil-Rhizosphere Stem Elongation</b>				
<i>Ambisporaceae</i>	1.126446	0.297184	1.528897	0.895664
<i>Chaetosphaeriaceae</i>	17.49084	-0.70025	1.581208	0.895664
<i>Cladosporiaceae</i>	499.0058	-0.33263	0.773925	0.895664
<i>Erythrobasidiaceae</i>	266.2979	2.846739	1.621902	0.369093
<i>Hypocreales Incertae sedis</i>	8407.459	-1.26002	0.733516	0.369093
<i>Mortierellaceae</i>	3348.64	3.034097	0.857525	0.017323
<i>Sporidiobolaceae</i>	979.9934	0.496755	0.954706	0.895664
<i>Sydowiellaceae</i>	810.7476	-1.01088	0.795848	0.461721
<i>Parmeliaceae</i>	946.0702	0.948964	0.83265	0.497266
<b>Comparison – Rhizosphere-Endosphere Stem Elongation</b>				
<i>Ambisporaceae</i>	3.5	2.584955	1.466257	0.136337
<i>Chaetosphaeriaceae</i>	6.666667	-2.23704	1.439501	0.168246
<i>Cladosporiaceae</i>	211.6667	-3.02444	1.295549	0.053817
<i>Erythrobasidiaceae</i>	191.5	-5.43269	1.6628	0.004345
<i>Hypocreales Incertae sedis</i>	2356.833	-2.30279	0.561104	0.000314
<i>Mortierellaceae</i>	2594.5	-6.9257	0.786027	3.47E-17
<i>Parmeliaceae</i>	2737.5	2.135562	0.52331	0.000314
<i>Sporidiobolaceae</i>	525.3333	-6.51894	1.131744	1.18E-07
<i>Sydowiellaceae</i>	239.1667	-3.20945	0.839198	0.000734

\*lfsSE= log2 Fold Change Standard Error, Lfc = log2 Fold Change

**Supplementary Table S.7. DESeq2 outputs, significantly differentially abundant fungal taxa**

Differential abundance analysis comparing the abundance of fungal taxa within the endosphere of senesced plants compared to stem elongation growth phase plants, or comparing the abundance of fungal taxa across compartments; compartment comparisons use pooled data for pot cultivated plants (Levington compost, agricultural soil, and the 50:50 mix) and field cultivated stem elongation growth phase plants, and compared abundances within the bulk soil to the rhizosphere, and the rhizosphere to the endosphere.

<b>Comparison - Endosphere Senescent-Endosphere Stem Elongation</b>				
<b>Taxon</b>	<b>Base Mean</b>	<b>Lfc</b>	<b>lfcSE*</b>	<b>padj</b>
<i>Parmeliaceae</i>	2232.5	-9.80089	0.980524	4.94E-22
<i>Leotiaceae</i>	2225.5	-12.1196	1.281843	5.02E-20
<i>Myxotrichaceae</i>	524.5	-7.1502	1.058346	1.47E-10
<i>Clavicipitaceae</i>	4820.833333	-3.95072	1.096925	0.0014004
<i>Chaetosphaeriaceae</i>	26.66666667	4.450025	1.070374	0.0001663
<b>Comparison – Bulk Soil-Rhizosphere all samples (compost, agricultural pot, 50:50 mix and field stem elongation)</b>				
<b>Taxon</b>	<b>Base Mean</b>	<b>Lfc</b>	<b>lfcSE*</b>	<b>padj</b>
<i>Australiascaceae</i>	736.0431973	-0.22788	0.803202	0.8926117
<i>Glomerellaceae</i>	2480.256112	-0.23093	0.96327	0.8926117
<i>Hypocreales Incertae sedis</i>	6355.970558	-1.04999	0.731854	0.509278
<i>Leotiaceae</i>	1357.387391	-1.72508	0.760808	0.2009295
<i>Mortierellaceae</i>	1382.193766	3.010589	0.87225	0.0119848
<b>Comparison – Rhizosphere-Endosphere all samples (compost, agricultural pot, 50:50 mix and field stem elongation)</b>				
<i>Australiascaceae</i>	616.0318099	-3.69862	0.722446	6.58E-06
<i>Glomerellaceae</i>	2071.880172	-3.3994	0.928719	0.0009847
<i>Hypocreales Incertae sedis</i>	3357.081144	-3.56054	0.733033	1.59E-05
<i>Leotiaceae</i>	1491.808491	3.60561	0.749014	1.59E-05
<i>Mortierellaceae</i>	1463.011582	-3.95723	1.01644	0.0004576
*lfsSE= log2 Fold Change Standard Error, Lfc = log2 Fold Change				

**Supplementary Table S.8. DESeq2 outputs to identify root exudate utilisers in the rhizosphere**

Differential abundance analysis comparing the abundance of bacterial taxa within the <sup>13</sup>C heavy fractions to the <sup>12</sup>C heavy, or <sup>13</sup>C light fractions. Bacterial taxa which were significantly greater in abundance within the <sup>13</sup>C heavy fraction for both comparisons were considered to be utilising root exudates.

<b><sup>12</sup>C heavy compared to <sup>13</sup>C heavy</b>				
<b>Taxon</b>	<b>baseMean</b>	<b>log2FoldChange</b>	<b>lfsSE *</b>	<b>Padj</b>
<i>Enterobacteriaceae</i>	1525.585247	9.225478608	1.330868073	8.30E-11
<i>Paenibacillaceae</i>	542.7147998	6.131222282	0.71096459	1.62E-16
<i>Verrucomicrobiaceae</i>	1540.476332	5.525652719	0.38986191	1.34E-43
<i>Pseudomonadaceae</i>	747.4943245	5.296358431	1.327440714	0.000347898
<i>Oxalobacteraceae</i>	1121.798139	5.093564169	0.466396414	3.05E-26
<i>Cellvibrionaceae</i>	143.5596665	4.563865623	0.692665528	6.33E-10
<i>Comamonadaceae</i>	1965.095913	4.553967836	0.399631593	2.20E-28
<i>Fibrobacteraceae</i>	308.4266565	4.250982079	0.691291828	9.73E-09
<i>Rhizobiaceae</i>	200.6141532	3.586599358	0.543771655	6.33E-10
<i>Cytophagaceae</i>	333.3918811	3.485162531	0.695142566	4.45E-06
<i>Micrococcaceae</i>	1002.919707	3.188857943	0.72279587	6.41E-05
<i>Microbacteriaceae</i>	182.5423665	2.316546635	0.708211293	0.004465608
<i>Xanthomonadaceae</i>	191.2474009	2.210211572	0.38114642	6.68E-08
<i>Intrasporangiaceae</i>	196.5957779	1.679541697	0.522170126	0.005191223
<i>Polyangiaceae</i>	174.8113772	1.441127742	0.521692567	0.016393349
<b><sup>13</sup>C light compared to <sup>13</sup>C heavy</b>				
<b>Taxon</b>	<b>baseMean</b>	<b>log2FoldChange</b>	<b>lfsSE *</b>	<b>Padj</b>
<i>Enterobacteriaceae</i>	1193.557819	5.663078218	1.200852813	1.47E-05
<i>Paenibacillaceae</i>	474.7391477	3.450277252	0.717541055	1.06E-05
<i>Verrucomicrobiaceae</i>	1240.232555	4.817514423	0.470756764	4.58E-23
<i>Pseudomonadaceae</i>	640.932411	2.82797024	1.152315253	0.034597109
<i>Oxalobacteraceae</i>	882.3908862	6.866216698	0.630972123	6.88E-26
<i>Cellvibrionaceae</i>	122.6921288	3.359934177	0.602489726	3.00E-07
<i>Comamonadaceae</i>	1557.064356	4.810895677	0.375167998	1.19E-35
<i>Fibrobacteraceae</i>	247.2869102	3.558583564	0.645095913	3.77E-07
<i>Rhizobiaceae</i>	194.8884915	1.676668472	0.519657919	0.004548681
<i>Cytophagaceae</i>	283.9658689	2.520107112	0.595381826	0.000113113
<i>Micrococcaceae</i>	918.575397	1.673919306	0.649494375	0.025268353
<i>Microbacteriaceae</i>	144.683441	2.15700887	0.693944173	0.006585131
<i>Xanthomonadaceae</i>	167.3075238	1.56520526	0.53711052	0.011040717
<i>Intrasporangiaceae</i>	158.6061631	1.50689272	0.517691543	0.011040717
<i>Polyangiaceae</i>	119.7725455	2.424270358	0.530116698	2.62E-05
*lfsSE= log2 Fold Change Standard Error				

Table S.9. Identity matrix showing the percent sequence identity for each *Burkholderiaceae* sequence recovered from the endosphere or rhizosphere of wheat, compared with isolates from this family and with representative sequences acquired from the NCBI database. Sequences labelled with \_cXXX, where XXX is a number which shows the number the number of reads for this unique 16S rRNA gene sequence that were recovered from the endosphere. R\_ denotes sequences recovered from the rhizosphere. All other sequences are either labelled as isolates, isolated by this work, or were sourced from the NCBI database. Each column and row is numbered for each unique sequence.

1	Herbaspirillum_c010	87.47	100	97.65	94.69	95.18	95.38	95.31	95.04	95.11	95.18	92.6	94.68	94.28	94.97	94.69	90.64	90.43	90.19	89.27	87.81	89.6	90.09	87.42	89.14	87.41	91.12	91.05
2	Herbaspirillum_robiniae	87.54	97.65	100	94.61	95.07	95.29	95.22	95.07	94.93	95.07	93.2	94.32	93.86	94.72	94.29	89.59	89.45	89.19	89.04	87.44	88.82	89.09	87.21	88.64	87.24	89.59	90.09
3	Duganella_phylosphaerae	87.15	94.69	94.61	100	98.38	98.38	98.45	98.23	98.23	98.38	94.79	97.05	96.61	97.05	97.05	89.45	89.15	89.19	89.06	88.05	88.86	89	88.11	88.73	88.03	89.24	90.04
4	Duganella_c242	87.85	95.18	95.07	98.38	100	99.72	99.66	99.59	99.72	99.66	94.94	96.83	96.69	97.24	97.11	90.01	89.66	89.7	89.19	88.54	89.58	89.8	88.45	89.5	88.24	90.14	90.9
5	Duganella_c353	87.78	95.38	95.29	98.38	99.72	100	99.79	99.59	99.59	99.66	94.94	96.83	96.55	97.11	96.97	89.94	89.73	89.63	88.98	88.24	89.38	89.73	88.15	89.31	87.89	90.01	90.7
6	R_Duganella	87.78	95.38	95.29	98.38	99.72	100	99.79	99.59	99.59	99.66	94.94	96.83	96.55	97.11	96.97	89.94	89.73	89.63	88.98	88.24	89.38	89.73	88.15	89.31	87.89	90.01	90.7
7	Duganella_c192	87.85	95.31	95.22	98.45	99.66	99.79	100	99.52	99.52	99.72	94.81	96.76	96.49	97.04	96.9	89.87	89.66	89.7	89.05	88.39	89.44	89.8	88.3	89.4	88.06	89.94	90.63
8	Duganella_c512	87.91	95.04	95.07	98.23	99.59	99.59	99.52	100	99.72	99.66	94.94	96.69	96.42	96.97	96.83	89.8	89.59	89.5	88.98	88.54	89.38	89.59	88.45	89.31	88.24	89.94	90.7
9	Duganella_c080	87.98	95.11	94.93	98.23	99.72	99.59	99.52	99.72	100	99.79	94.94	96.62	96.49	97.04	96.9	89.87	89.52	89.56	89.05	88.54	89.44	89.66	88.45	89.31	88.24	90.01	90.76
10	Duganella_c526	88.12	95.18	95.07	98.38	99.66	99.66	99.72	99.66	99.79	100	94.94	96.62	96.49	97.04	96.9	89.87	89.52	89.7	89.05	88.54	89.44	89.8	88.45	89.31	88.24	89.94	90.63
11	Massilia_Isolate_RNA126	87.65	92.6	93.2	94.79	94.94	94.94	94.81	94.94	94.94	94.94	100	97.12	97.66	97.27	97.4	88.7	88.44	87.86	87.86	88.42	88.44	88.44	88.13	88.8	88.21	88.96	89.74
12	Massilia	87.76	94.68	94.32	97.05	96.83	96.83	96.76	96.69	96.62	96.62	97.12	100	99.14	99.42	99.57	89.14	88.92	88.89	88.61	87.52	88.78	88.85	87.29	88.78	87.37	89.58	89.64
13	R_Massilia_violaceinigra_R_A	88.26	94.28	93.86	96.61	96.69	96.55	96.49	96.42	96.49	96.49	97.66	99.14	100	99.17	99.59	89.39	89.04	89.15	88.35	88.08	89.17	89.39	87.69	89.03	87.72	90.01	89.87
14	Massilia_violaceinigra_c945	87.85	94.97	94.72	97.05	97.24	97.11	97.04	96.97	97.04	97.04	97.27	99.42	99.17	100	99.59	89.66	89.32	89.29	88.49	87.62	89.03	89.25	87.23	88.74	87.2	90.14	90.14
15	R_Massilia_violaceinigra_R_B	88.12	94.69	94.29	97.05	97.11	96.97	96.9	96.9	96.9	97.4	99.57	99.59	99.59	100	89.52	89.18	89.29	88.49	87.77	89.03	89.25	87.39	88.84	87.37	90.01	90.01	
16	R_Rhizobacter	86.28	90.64	89.59	89.45	90.01	89.94	89.87	89.8	89.87	89.87	88.7	89.14	89.39	89.66	89.52	100	99.52	93.09	92.23	91.2	92.23	93.05	91.52	91.78	90.52	93.6	93.87
17	Rhizobacter_fulvus	85.93	90.43	89.45	89.15	89.66	89.73	89.66	89.59	89.52	89.52	88.44	88.92	89.04	89.32	89.18	99.52	100	92.74	91.95	90.59	91.98	92.7	90.91	91.41	89.83	93.26	93.76
18	Polaromonas_c028	85.77	90.19	89.19	89.19	89.7	89.63	89.7	89.5	89.56	89.7	87.86	88.89	89.15	89.29	89.29	93.09	92.74	100	98.45	93.93	95.83	96.35	94.19	95.16	93.55	95.52	96.62
19	Polaromonas_eurypsychrophila	84.98	89.27	89.04	89.06	89.19	88.98	89.05	88.98	89.05	89.05	87.86	88.61	88.35	88.49	88.49	92.23	91.95	98.45	100	93.14	95.28	95.56	93.43	94.68	92.68	95.29	96.62
20	Variovorax_Isolate_RR205	86.4	87.81	87.44	88.05	88.54	88.24	88.39	88.54	88.54	88.54	88.42	87.52	88.08	87.62	87.77	91.2	90.59	93.93	93.14	100	98.15	98.46	98.9	98.75	99.14	94.75	94.75
21	Variovorax_paradoxus	86.63	89.6	88.82	88.86	89.58	89.38	89.44	89.38	89.44	89.44	88.44	88.78	89.17	89.03	89.03	92.23	91.98	95.83	95.28	98.15	100	99.03	99.24	99.43	99.31	96.4	96.78
22	R_Variovorax_paradoxus	86.86	90.09	89.09	89	89.8	89.73	89.8	89.59	89.66	89.8	88.44	88.85	89.39	89.25	89.25	93.05	92.7	96.35	95.56	98.46	99.03	100	99.55	99.34	99.48	96.84	97.32
23	Variovorax_Isolate_RR302	86.19	87.42	87.21	88.11	88.45	88.15	88.3	88.45	88.45	88.45	88.13	87.29	87.69	87.23	87.39	91.52	90.91	94.19	93.43	98.9	99.24	99.55	100	99.85	100	94.39	94.55
24	Variovorax_Isolate_RR307	86.67	89.14	88.64	88.73	89.5	89.31	89.4	89.31	89.31	89.31	88.8	88.78	89.03	88.74	88.84	91.78	91.41	95.16	94.68	98.75	99.43	99.34	99.85	100	100	95.66	96.22
25	Variovorax_Isolate_RR304	86.36	87.41	87.24	88.03	88.24	87.89	88.06	88.24	88.24	88.24	88.21	87.37	87.72	87.2	87.37	90.52	89.83	93.55	92.68	99.14	99.31	99.48	100	100	100	93.79	93.97
26	Rhodoferax_c294	86.38	91.12	89.59	89.24	90.14	90.01	89.94	89.94	90.01	89.94	88.96	89.58	90.01	90.14	90.01	93.6	93.26	95.52	95.29	94.75	96.4	96.84	94.39	95.66	93.79	100	98.15
27	Rhodoferax_sediminis	86.18	91.05	90.09	90.04	90.9	90.7	90.63	90.7	90.76	90.63	89.74	89.64	89.87	90.14	90.01	93.87	93.76	96.62	96.62	94.75	96.78	97.32	94.55	96.22	93.97	98.15	100
	<b>Sequences</b>	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27

Table S.10. Identity matrix showing the percent sequence identity for each *Streptomyces* sequence recovered from the endosphere of wheat, compared with isolates from this family and with representative sequences acquired from the NCBI database. PC indicates a sequence was acquired from PacBio sequencing. Endosphere sequences are labelled with \_cXXX, where XXX is a number which shows the number of reads for this unique 16S rRNA gene sequence recovered from the endosphere. All rows labelled with with a letter/number format (such as CRS3 or CRwSp2) refer to different *Streptomyces* endosphere isolates, and other sequences are representatives for different species acquired from the NCBI database.

PB_Streptomyces_turgidis_cabies_c275	PB_Streptomyces_turgidis_cabies_c684	PB_Streptomyces_canus_c068	PB_Streptomyces_turgidis_cabies_c161	PB_Streptomyces_turgidis_cabies_c113	PB_Streptomyces_scabiei_c025	Sequences
98	98	98	97	98	97	PRS1
98	98	98	97	97	97	CRwS3
98	98	98	97	97	97	SRwS1
98	98	98	97	98	97	PRwS2
98	98	98	97	98	97	CRS5
98	98	98	97	97	97	PRS3
98	98	98	97	98	97	CRS1
98	98	98	98	98	97	SRS4
98	98	98	98	98	97	SRS5
98	98	98	98	98	97	CRS2
98	97	97	97	97	96	PEI5
98	97	97	97	97	96	PES3
98	97	97	97	97	96	PEI6
98	98	97	97	97	97	SRS3
98	98	97	97	97	97	CESp1
98	98	97	98	98	97	CEIp1
97	98	98	97	97	96	SRwS2
97	98	98	97	97	96	PRwS1
98	98	98	97	97	97	PES5
97	98	98	97	97	97	SRwSp2
97	98	98	97	97	97	SES5
97	97	97	97	97	96	CRwS4
98	98	98	98	98	97	CRwSp2
97	97	98	97	97	97	PRS2
98	99	98	98	98	97	SES3
98	98	98	98	98	97	PES4
98	98	98	98	98	97	SRS1
100	99	98	99	99	97	PB_Streptomyces_turgidis_cabies_c275
99	100	97	99	99	97	PB_Streptomyces_turgidis_cabies_c684
99	100	97	99	99	97	CESp2
99	100	97	99	99	97	CESp3
98	97	100	97	97	98	Streptomyces_canus
98	97	100	97	97	98	PB_Streptomyces_canus_c068
97	97	100	97	97	98	CRwSp2b
97	97	100	97	97	98	SRwSp1
97	97	100	97	97	98	PRwSp2
97	97	99	97	97	98	CRwSp1
97	97	99	97	97	98	PRwSp1
97	97	99	97	97	98	PRwS3
98	97	100	97	97	98	PES2
97	97	99	97	97	98	PRwlp3
99	99	97	99	99	98	Streptomyces_turgidiscabie
98	98	98	98	98	98	CRwS1
98	98	98	98	98	98	PRwS4
98	98	98	98	98	98	SRwS3
98	98	98	98	98	98	PES1
98	98	98	98	98	98	CRS4
99	99	97	100	99	97	PB_Streptomyces_turgidis_cabies_c161
98	98	98	99	98	98	CRwS5
99	99	97	99	100	97	PB_Streptomyces_turgidis_cabies_c113
97	97	99	97	97	98	CRwS2
96	97	97	97	97	96	PRwlp1
96	97	97	97	97	96	PRwS5
96	97	97	97	97	96	CRS3
98	98	98	98	98	98	PRS5
97	97	98	97	97	97	PRwlp2
97	97	98	97	97	97	SES1
97	97	99	97	97	98	SRS2
96	96	96	96	96	96	Streptomyces_lydicus
97	97	98	97	97	100	PB_Streptomyces_scabiei_c025
96	96	98	96	96	96	Streptomyces_coelicolor
96	96	98	96	96	96	SRwS4
96	96	96	96	96	95	Streptomyces_griseoviridis
89	89	89	89	89	89	SES4
89	89	89	89	89	89	SES2
96	96	95	96	96	95	Streptomyces_venezuelae_NRLL_B-65442
94	94	97	94	94	99	Streptomyces_scabiei_87_22

Table S.11. Identity matrix showing the percent sequence identity for each *Pseudomonadaceae* sequence recovered from the endosphere or rhizosphere of wheat, compared with isolates from this family and with representative sequences acquired from the NCBI database. Endosphere sequences are labelled with \_cXXX, where XXX is a number which shows the number of reads for this unique 16S rRNA gene sequence recovered from the endosphere. R\_ denotes rhizosphere derived sequences. All other sequences are either labelled as isolates, isolated by this work, or were sourced from the NCBI database.

Pseudomonas_c191	100	99	99	99	98	98	98	97	97
R_Pseudomonas_poae	99	100	99	99	98	98	98	97	98
Pseudomonas_brassicacearum_Isolate_MNA132	98	98	98	100	99	100	99	99	99
Pseudomonas_brassicacearum_Isolate_RNL304	98	98	98	99	100	99	99	99	99
Pseudomonas_Isolate_RNL309	97	99	98	99	99	99	99	99	100
Pseudomonas_Isolate_RNL311	97	98	97	99	99	99	99	100	99
Pseudomonas_Isolate_RR104	98	98	98	99	99	99	100	99	99
<b>Sequences</b>	Pseudomonas_c191	Pseudomonas_poae	R_Pseudomonas_poae	Pseudomonas_brassicacearum	Pseudomonas_brassicacearum_Isolate_RNL304	Pseudomonas_brassicacearum_Isolate_MNA132	Pseudomonas_Isolate_RR104	Pseudomonas_Isolate_RNL311	Pseudomonas_Isolate_RNL309

Table S.12. Identity matrix showing the percent sequence identity for each *Rhizobiaceae* sequence recovered from the endosphere of wheat, compared with representative sequences acquired from the NCBI database. Endosphere sequences are labelled with \_cXXX, where XXX is a number which shows the number of reads for this unique 16S rRNA gene sequence recovered from the endosphere. All other sequences were sourced from the NCBI database.

Rhizobiaceae_c033	100	94.38	93.53	94.46	95.07	94.5	95.17	95.17
Mesorhizobium_huakuii	94.38	100	95.73	96.37	95.14	94.77	94.24	95.02
Phyllobacterium_c13	93.53	95.73	100	97.16	93.28	92.49	93.32	94.74
Phyllobacterium_phragmitis	94.46	96.37	97.16	100	94.42	93.78	94.1	95.66
Neorhizobium_galegae_c07	95.07	95.14	93.28	94.42	100	98.93	95.93	96.07
Neorhizobium_galegae	94.5	94.77	92.49	93.78	98.93	100	95.28	95.5
Rhizobiaceae_c093	95.17	94.24	93.32	94.1	95.93	95.28	100	96.09
Rhizobiaceae_c085	95.17	95.02	94.74	95.66	96.07	95.5	96.09	100
<b>Sequence</b>	Rhizobiaceae_c033	Mesorhizobium_huakuii	Phyllobacterium_c13	Phyllobacterium_phragmitis	Neorhizobium_galegae_c07	Neorhizobium_galegae	Rhizobiaceae_c093	Rhizobiaceae_c085

Table S.13. Identity matrix showing the percent sequence identity for each *Chitinophagaceae* sequence recovered from the endosphere or rhizosphere of wheat, compared with representative sequences acquired from the NCBI database. Endosphere sequences are labelled with \_cXXX, where XXX is a number which shows the number of reads for this unique 16S rRNA gene sequence recovered from the endosphere. R\_ denotes rhizosphere sequences. All other sequences were sourced from the NCBI database.

Lewinella_lacunae	100	81.06	80.78	80.65	80.25	80.38	80.6	80.07
Chitinophaga_c58	81.06	100	97.1	90.89	89.86	89.79	89.86	89.52
Chitinophaga_silvisoli	80.78	97.1	100	91.04	90.12	90.06	89.92	90.42
Niastella_hibisci	80.65	90.89	91.04	100	95.73	94.42	92.91	92.34
Niastella_c92	80.25	89.86	80.12	95.73	100	96.42	91.95	91.8
Niastella_c56	80.38	89.79	80.06	94.42	96.42	100	90.9	92.07
R_Terrimonas	80.6	89.86	89.92	92.91	91.95	90.9	100	95.04
Terrimonas_ferruginea	80.07	89.52	80.42	92.34	91.8	92.07	95.04	100
<b>Sequences</b>	Lewinella_lacunae	Chitinophaga_c58	Chitinophaga_silvisoli	Niastella_hibisci	Niastella_c92	Niastella_c56	R_Terrimonas	Terrimonas_ferruginea

Table S.14. Identity matrix showing the percent sequence identity for each *Micrococcaceae* sequence recovered from the endosphere or rhizosphere of wheat, compared with representative sequences acquired from the NCBI database. Endosphere sequences are labelled with \_cXXX, where XXX is a number which shows the number of reads for this unique 16S rRNA gene sequence recovered from the endosphere. R\_ denotes rhizosphere sequences. All other sequences were sourced from the NCBI database.

Bifidobacterium_longum	100	85.86	85.87	86.34	85.3	85.58	84.93	85.43
Arthrobacter_humicola_c40	85.86	100	99.86	98.62	97.72	97.86	97.63	98.2
Arthrobacter_humicola	85.87	99.86	100	98.69	97.78	97.98	97.69	98.33
Arthrobacter_pascens	86.34	98.62	98.69	100	98.75	98.82	97.72	98.27
Arthrobacter_pascens_c86	85.3	97.72	97.78	98.75	100	99.65	97.49	97.51
R_Arthrobacter_pascens	85.58	97.86	97.98	98.82	99.65	100	97.14	97.85
Pseudarthrobacter_sulfonivorans	84.93	97.63	97.69	97.72	97.49	97.14	100	98.74
R_Pseudarthrobacter_sulfonivorans	85.43	98.2	98.33	98.27	97.51	97.85	98.74	100
<b>Sequences</b>	Bifidobacterium_longum	Arthrobacter_humicola_c40	Arthrobacter_humicola	Arthrobacter_pascens	Arthrobacter_pascens_c86	R_Arthrobacter_pascens	Pseudarthrobacter_sulfonivorans	R_Pseudarthrobacter_sulfonivorans

Table S.15. Identity matrix showing the percent sequence identity for each *Paenibacillaceae* sequence recovered from the rhizosphere of wheat, compared with representative sequences acquired from the NCBI database. R\_ denotes rhizosphere sequences. All other sequences were sourced from the NCBI database.

R_Cohnella	100	96.14	90.43
Cohnella_endophytica	96.14	100	91.45
Paenibacillus_zeisoli	90.43	91.45	100
<b>Sequences</b>	R_Cohnella	Cohnella_endophytica	Paenibacillus_zeisoli

Table S.15. Identity matrix showing the percent sequence identity for each *Cytophagaceae* sequence recovered from the endosphere or rhizosphere of wheat, compared with representative sequences acquired from the NCBI database. Endosphere sequences are labelled with \_cXXX, where XXX is a number which shows the number of reads for this unique 16S rRNA gene sequence recovered from the endosphere. All other sequences were sourced from the NCBI database.

Sporocytophaga_myxococcoides	100	87.39	87.94
Cytophaga_hutchinsonii_c8	87.39	100	98.54
Cytophaga_hutchinsonii	87.94	98.54	100
<b>Sequences</b>	Sporocytophaga_myxococcoides	Cytophaga_hutchinsonii_c8	Cytophaga_hutchinsonii